A sample-paths approach to noise-induced synchronization: Stochastic resonance in a double-well potential

Nils Berglund and Barbara Gentz

Abstract

Additive white noise may significantly increase the response of bistable systems to a periodic driving signal. We consider two classes of double-well potentials, symmetric and asymmetric, modulated periodically in time with period $1/\varepsilon$, where ε is a moderately (not exponentially) small parameter. We show that the response of the system changes drastically when the noise intensity σ crosses a threshold value. Below the threshold, paths are concentrated near one potential well, and have an exponentially small probability to jump to the other well. Above the threshold, transitions between the wells occur with probability exponentially close to $1/2$ in the symmetric case, and exponentially close to 1 in the asymmetric case. The transition zones are localised in time near the points of minimal barrier height. We give a mathematically rigorous description of the behaviour of individual paths, which allows us, in particular, to determine the power-law dependence of the critical noise intensity on ε and on the minimal barrier height, as well as the asymptotics of the transition and non-transition probabilities.

Date. December 29, 2000.

2000 *Mathematics Subject Classification.* 37H99, 60H10 (primary), 34F05, 34E15 (secondary).

Keywords and phrases. Stochastic resonance, noise-induced synchronization, double-well potential, additive noise, random dynamical systems, non-autonomous stochastic differential equations, singular perturbations, pathwise description, concentration of measure.

1 Introduction

Since its introduction as a model for the periodic appearance of the ice ages [BPSV], stochastic resonance has been observed in a large number of physical and biological systems, including lasers, electronic circuits and the sensory system of crayfish (for reviews of applications, see for instance [MW]).

The mechanism of stochastic resonance can be illustrated in a simple model. Consider the overdamped motion of a particle in a double-well potential. The two potential wells describe two macroscopically different states of the unperturbed system, for instance cold and warm climate. The particle is subject to two different kinds of perturbation: a deterministic periodic driving force (such as the periodic variation of insulation caused by the changing eccentricity of the earth's orbit), and an additive noise (modeling the random influence of the weather). Each of these two perturbations, taken by itself, does not produce any interesting dynamics (from the point of view of resonance). Indeed, the periodic driving is assumed to have too small an amplitude to allow for any transitions between the potential wells in the absence of noise. On the other hand, without periodic forcing, additive noise will cause the particle to jump from one potential well to the other

at random times. The expected time between transitions is given asymptotically, in the small noise limit, by Kramers' time, which is proportional to the exponential of the barrier height H over the noise intensity squared, namely e^{H/σ^2}. When both perturbations are combined, however, and their amplitudes suitably tuned, the particle will flip back and forth between the wells in a close to periodic way. Thus the internal noise can significantly enhance the weak external periodic forcing, by producing large amplitude oscillations of the system, hence the name of resonance.

The choice of the term "resonance" has been questioned, as "it would be more appropriate to refer to *noise-induced signal-to-noise ratio enhancement*" [Fox]. In the regime of a periodic driving whose amplitude is not a small parameter, one also speaks of *noise-induced synchronization* [SNAS]. Appreciable, though still sub-threshold amplitudes of the periodic driving have the advantage to enable transitions for small noise intensities, without requiring astronomically long driving periods.

While the heuristic mechanism of stochastic resonance is rather well understood, a complete mathematical description is still lacking, though important progress has been made in several limiting cases. Depending on the regime one is interested in, several approaches have been used to describe the phenomenon quantitatively. The simplest ones use a discretization of either time or space. When the potential is considered as piecewise constant in time, the generator of the autonomous case can be used to give a complete solution [BPSV], showing that resonance occurs when driving period and Kramers' time are equal. Alternatively, space can be discretized in order to obtain a two-state model, which is described by a Markovian jump process [ET]. The two-state model has also been realised experimentally by an electronic circuit, called the Schmitt trigger [FH, McNW].

In physical experiments, one has often access to indirect characteristics of the dynamics, such as the power spectrum, which displays a peak at the driving frequency. The strength of the resonance is quantified through the signal-to-noise ratio (SNR), which is proportional to the area under the peak (this definition obviously leaves some liberty of choice). The SNR has been estimated, in the limit of small driving amplitude, by using spectral theory of the Fokker–Planck equation [Fox, JH], or a "rate" equation for the probability density [McNW]. The signal-to-noise ratio is found to behave like $\mathrm{e}^{-H/\sigma^2}/\sigma^4$, which reaches a maximum for $\sigma^2 = H/2$.

The probability density of the process, however, only gives part of the picture, and a more detailed understanding of the behaviour of individual paths is desirable. Some interesting progress in this direction is found in [Fr]. The approach applies to a very general class of dynamical systems, in the limit of vanishing noise intensity. When the period of the forcing scales like Kramers' time, solutions of the stochastic differential equation are shown to converge to periodic functions in the following sense: The L^p-distance between the paths and the periodic limiting function converges to zero in probability as the noise intensity goes to zero. Due to its generality, however, this approach does not give any information on the rate of convergence of typical paths to the periodic function, nor does it estimate the probability of atypical paths. Also, since the period of the forcing must scale like Kramers' time, the assumed small noise intensity goes hand in hand with exponentially long waiting times between interwell transitions.

In the present work, we provide a more detailed description of the individual paths' behaviour, for small but finite noise intensities and driving frequencies. We consider two classes of one-dimensional double-well potentials, symmetric and asymmetric ones. The height of the potential barrier is assumed to become small periodically, which allows us to

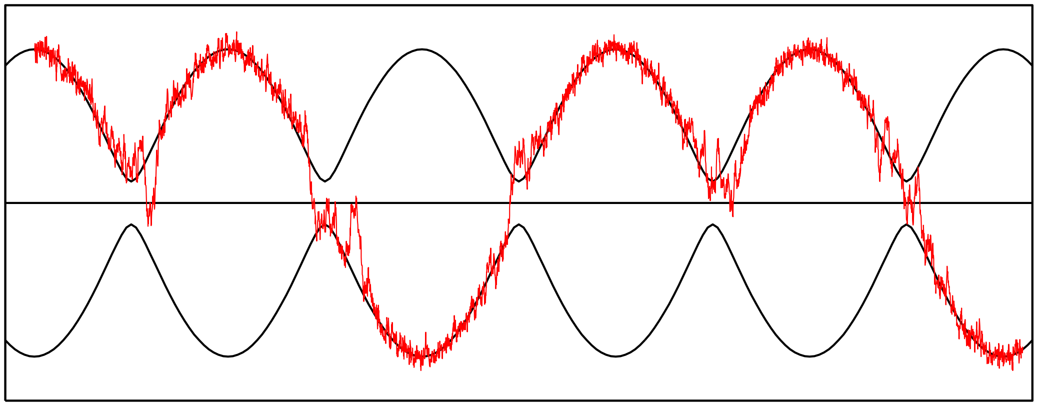

FIGURE 1. A typical solution of the SDE (1.1) in the case of the symmetric potential (1.2). Heavy curves indicate the position of the potential wells, which approach each other at integer times. The straight line is the location of the saddle. Parameter values $\varepsilon = 0.01$, $\sigma = 0.08$ and $a_0 = 0.02$ belong to the regime where the transition probability between wells is close to $1/2$. We show that transitions are concentrated in regions of order $\sqrt{\sigma}$ around the instants of minimal barrier height.

consider situations where the period need not be exponentially large in $1/\sigma^2$ for transitions between the wells to be likely.

In the case of an asymmetric potential, we are interested, in particular, in determining the optimal noise intensity as a function of the driving frequency and the minimal barrier height, guaranteeing a close-to-periodic oscillation between both wells. We will estimate both the deviation (in space and time) of typical paths from the limiting periodic function, and the asymptotics of the probability of exceptional paths. The case of a symmetric potential shows an additional feature. For the right choice of the noise intensity, transitions become likely once per period, at which time the "new" well is chosen at random. We will again estimate the deviation from a suitable reference process and the asymptotics of the probability of exceptional paths.

The systems are described by stochastic differential equations (SDEs) of the form

$$\mathrm{d}x_s = -\frac{\partial}{\partial x} V(x_s, s)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s, \tag{1.1}$$

where W_s is a Brownian motion. The potential $V(x,s)$ is $1/\varepsilon$-periodic in s, and admits two minima for every value of s. The frequency ε, the minimal barrier height between the wells and the noise intensity σ are considered as (moderately) small parameters, the relation between which will determine the transition probability.

The first class of potentials we consider is symmetric in x. A typical representative of this class is the potential

$$V(x,s) = -\frac{1}{2}a(\varepsilon s)x^2 + \frac{1}{4}x^4, \qquad \text{with } a(\varepsilon s) = a_0 + 1 - \cos(2\pi\varepsilon s). \tag{1.2}$$

Here $a_0 \geqslant 0$ is a parameter controlling the minimal barrier height. We introduce the slow time $t = \varepsilon s$ for convenience. The potential has two wells, located at $\pm\sqrt{a(t)}$, separated by a barrier of height $\frac{1}{4}a(t)^2$. The distance between the wells and the barrier height become small simultaneously, at integer values of t.

Our results for symmetric potentials can be summarized as follows:

- In the deterministic case $\sigma = 0$, we describe the dependence of solutions on t, a_0 and ε (Theorem 2.1). Solutions starting at $x > 0$ are attracted by the potential well at $\sqrt{a(t)}$, which they track with a small lag. If $a_0 \geqslant \varepsilon^{2/3}$, this lag is at most of the order ε/a_0; if $a_0 \leqslant \varepsilon^{2/3}$, it is at most of the order $\varepsilon^{1/3}$, but solutions never approach the saddle closer than a distance of order $\varepsilon^{1/3}$ (even if $a_0 = 0$).
- When noise is present, but σ is small compared to the maximum of a_0 and $\varepsilon^{2/3}$, the paths are likely to track the solution of the corresponding deterministic differential equation at a distance of order $\sigma/\max\{|t|, \sqrt{a_0}, \varepsilon^{1/3}\}$ (Theorem 2.2). The probability to reach the saddle during one time period is exponentially small in $\sigma^2/(\max\{a_0, \varepsilon^{2/3}\})^2$.
- If σ is larger than both a_0 and $\varepsilon^{2/3}$, transitions between potential wells become likely, but are concentrated on the time interval $[-\sqrt{\sigma}, \sqrt{\sigma}]$ (repeated periodically). During this time interval, the paths may jump back and forth frequently between both potential wells, and they have a typical spreading of the order $\sigma/\max\{\sqrt{a_0}, \varepsilon^{1/3}\}$. After time $\sqrt{\sigma}$, the paths are likely to choose one of the wells and stay there till the next period (Theorem 2.4). The probability to choose either potential well is exponentially close to 1/2, with an exponent of order $\sigma^{3/2}/\varepsilon$, which is independent of a_0 (Theorem 2.3).
- This picture remains true when σ is larger than both $\sqrt{a_0}$ and $\varepsilon^{1/3}$, but note that the spreading of paths during the transition may become very large. Thus increasing noise levels will gradually blur the periodic signal.

These results show a rather sharp transition to take place at $\sigma = \max\{a_0, \varepsilon^{2/3}\}$, from a regime where the paths are unlikely to switch from one potential well to the other one, to a regime where they do switch with a probability exponentially close to 1/2 (Fig. 1).

The second class of potentials we consider is asymmetric, a typical representative being

$$V(x, s) = -\frac{1}{2}x^2 + \frac{1}{4}x^4 - \lambda(\varepsilon s)x. \tag{1.3}$$

This is a double-well potential if and only if $|\lambda| < \lambda_c = 2/(3\sqrt{3})$. We thus choose $\lambda(\varepsilon s) = \lambda(t)$ of the form

$$\lambda(t) = -(\lambda_c - a_0)\cos(2\pi t). \tag{1.4}$$

Near $t = 0$, the right-hand potential well approaches the saddle at a distance of order $\sqrt{a_0}$, and the barrier height is of order $a_0^{3/2}$. A similar encounter between the left-hand potential well and the saddle occurs at $t = 1/2$.

Our results for asymmetric potentials can be summarized as follows:

- In the deterministic case, solutions track the potential wells at a distance at most of order $\min\{\varepsilon/a_0, \sqrt{\varepsilon}\}$. If $a_0 \leqslant \varepsilon$, they never approach the saddle closer than a distance of order $\sqrt{\varepsilon}$ (Theorem 2.5).
- When σ is small compared to the maximum of $a_0^{3/4}$ and $\varepsilon^{3/4}$, paths are likely to track the deterministic solutions at a distance of order $\sigma/\max\{\sqrt{|t|}, a_0^{1/4}, \varepsilon^{1/4}\}$ (Theorem 2.6). The probability to overcome the barrier is exponentially small in $\sigma^2/(\max\{a_0^{3/4}, \varepsilon^{3/4}\})^2$.
- For larger σ, transitions become probable during the time interval $[-\sigma^{2/3}, \sigma^{2/3}]$. Due to the asymmetry, the probability to jump from the less deep potential well to the deeper one is exponentially close to one, with an exponent of order $\sigma^{4/3}/\varepsilon$, while paths are unlikely to come back (Theorem 2.7).

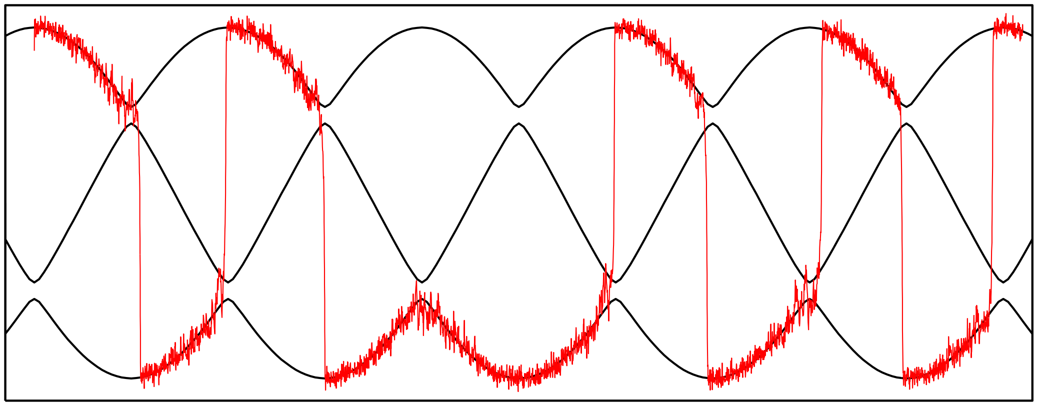

FIGURE 2. A typical solution of the SDE (1.1) in the case of the asymmetric potential (1.3). The upper and lower heavy curves indicate the position of the potential wells, while the middle curve is the location of the saddle. Parameter values $\varepsilon = 0.005$, $\sigma = 0.08$ and $a_0 = 0.005$ belong to the regime where the transition probability between wells is close to 1. We show that transitions are concentrated in regions of order $\sigma^{2/3}$ around the instants of minimal barrier height.

- This picture remains true when σ is larger than both $a_0^{1/4}$ and $\varepsilon^{1/4}$, but the spreading of paths during the transition may become very large.

Again, we find a rather sharp transition to take place, this time at $\sigma = \max\{a_0^{3/4}, \varepsilon^{3/4}\}$. In contrast to the symmetric case, for large σ the paths are likely to jump from one potential well to the other at every half-period (Fig. 2).

In both the symmetric and the asymmetric case, we thus obtain a high switching probability between the potential wells even for small noise intensities, provided minimal barrier height and driving frequency are sufficiently small. They only need, however, to be smaller than a power of σ: $a_0 \ll \sigma$ and $\varepsilon \ll \sigma^{3/2}$ in the symmetric case, and $a_0 \ll \sigma^{4/3}$, $\varepsilon \ll \sigma^{4/3}$ in the asymmetric case are sufficient conditions for switching dynamics.

Our results require a precise understanding of dynamical effects, and the subtle interplay between the probability to reach the potential barrier, the time needed for such excursions, and the total number of excursions with a chance of success. In this respect, they provide a substantial progress compared to the "quasistatic" approach, which considers potentials that are piecewise constant in time. Note that some of our results may come as a surprise. In particular, neither the width (in time) of the transition zone nor the asymptotics of the transition probability depend on the minimal barrier height a_0. In fact, the picture is independent of a_0 as soon as a_0 is smaller than $\varepsilon^{2/3}$ (in the symmetric case) or ε (in the asymmetric case), even for $a_0 = 0$. This is due to the fact that when a_0 is small, the time during which the potential barrier is low is too short to contribute significantly to the transition probability.

The remainder of this paper is organized as follows. The results are formulated in detail in Section 2, Subsection 2.2 being devoted to symmetric potentials, and Subsection 2.3 to asymmetric potentials. Section 3 contains the proofs for the symmetric case, while Section 4 contains the proofs for the asymmetric case.

Acknowledgements:

The fascinating problems related to stochastic resonance were brought to our attention by Anton Bovier, whom we would like to thank for stimulating discussions. N. B. thanks the

WIAS for kind hospitality.

2 Results

2.1 Preliminaries

We consider non-autonomous SDEs of the form (1.1). Introducing the slow time $t = \varepsilon s$ allows to study the system on a time interval of order one. When substituting t for εs, Brownian motion is rescaled and we obtain an SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon} f(x_t, t)\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad x_{t_0} = x_0, \tag{2.1}$$

where f is the force, derived from the potential V, and $\{W_t\}_{t\geqslant t_0}$ is a standard Wiener process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Initial conditions x_0 are always assumed to be square-integrable with respect to $\mathbb{P}$ and independent of $\{W_t\}_{t\geqslant t_0}$. Without further mentioning we always assume that f satisfies the usual (local) Lipschitz and bounded-growth conditions which guarantee existence and pathwise uniqueness of a strong solution $\{x_t\}_t$ of (2.1). Under these conditions, there exists a continuous version of $\{x_t\}_t$. Therefore we may assume that the paths $\omega \mapsto x_t(\omega)$ are continuous for $\mathbb{P}$-almost all $\omega \in \Omega$.

We introduce the notation $\mathbb{P}^{t_0,x_0}$ for the law of the process $\{x_t\}_{t\geqslant t_0}$, starting in x_0 at time t_0, and use $\mathbb{E}^{t_0,x_0}$ to denote expectations with respect to $\mathbb{P}^{t_0,x_0}$. Note that the stochastic process $\{x_t\}_{t\geqslant t_0}$ is an inhomogeneous Markov process. We are interested in first exit times of x_t from space–time sets. Let $\mathcal{A} \subset \mathbb{R} \times [t_0, t_1]$ be Borel-measurable. Assuming that $\mathcal{A}$ contains (x_0, t_0), we define the first exit time of (x_t, t) from $\mathcal{A}$ by

$$\tau_{\mathcal{A}} = \inf\bigl\{t \in [t_0, t_1] \colon (x_t, t) \notin \mathcal{A}\bigr\}, \tag{2.2}$$

and agree to set $\tau_{\mathcal{A}}(\omega) = \infty$ for those $\omega \in \Omega$ which satisfy $(x_t(\omega), t) \in \mathcal{A}$ for all $t \in [t_0, t_1]$. For convenience, we shall call $\tau_{\mathcal{A}}$ the *first exit time of x_t from $\mathcal{A}$*. Typically, we will consider sets of the form $\mathcal{A} = \{(x, t) \in \mathbb{R} \times [t_0, t_1] \colon g_1(t) < x < g_2(t)\}$ with continuous functions $g_1 < g_2$. Note that in this case, $\tau_{\mathcal{A}}$ is a stopping time[1] with respect to the canonical filtration of $(\Omega, \mathcal{F}, \mathbb{P})$ generated by $\{x_t\}_{t\geqslant t_0}$.

Before turning to the precise statements of our results, let us introduce some notations. We shall use

- $\lceil y \rceil$ for $y \geqslant 0$ to denote the smallest integer which is greater than or equal to y, and
- $y \vee z$ and $y \wedge z$ to denote the maximum or minimum, respectively, of two real numbers y and z.
- If $\varphi(t, \varepsilon)$ and $\psi(t, \varepsilon)$ are defined for small ε and for t in a given interval I, we write $\psi(t, \varepsilon) \asymp \varphi(t, \varepsilon)$ if there exist strictly positive constants $c_\pm$ such that $c_-\varphi(t, \varepsilon) \leqslant \psi(t, \varepsilon) \leqslant c_+\varphi(t, \varepsilon)$ for all $t \in I$ and all sufficiently small ε. The constants $c_\pm$ are understood to be independent of t and ε (and hence also independent of quantities like σ and a_0 which we consider as functions of ε).
- By $g(u) = \mathcal{O}(u)$ we indicate that there exist $\delta > 0$ and $K > 0$ such that $g(u) \leqslant Ku$ for all $u \in [0, \delta]$, where δ and K of course do not depend on ε or on the other small parameters a_0 and σ. Similarly, $g(u) = \mathcal{o}(1)$ is to be understood as $\lim_{u\to 0} g(u) = 0$.

Finally, let us point out that most estimates hold for small enough ε only, and often only for $\mathbb{P}$-almost all $\omega \in \Omega$. We will stress these facts only where confusion might arise.

[1] For a general Borel-measurable set $\mathcal{A}$, the first exit time $\tau_{\mathcal{A}}$ is still a stopping time with respect to the canonical filtration, completed by the null sets.

2.2 Symmetric case

We consider in this subsection the SDE (2.1) in the case of f being periodic in t, odd in x, and admitting two stable equilibrium branches, with a "barrier" between the branches becoming small once during every time period. A typical example of such a function is

$$f(x,t) = a(t)x - x^3 \qquad \text{with} \qquad a(t) = a_0 + 1 - \cos 2\pi t. \tag{2.3}$$

We will consider a more general class of functions $f : \mathbb{R}^2 \to \mathbb{R}$, which we assume to satisfy the following hypotheses:

- *Smoothness:* $f \in \mathcal{C}^4(\mathcal{M}, \mathbb{R})$, where $\mathcal{M} = [-d, d] \times \mathbb{R}$ and $d > 0$ is a constant;
- *Periodicity:* $f(x, t+1) = f(x,t)$ for all $(x,t) \in \mathcal{M}$;
- *Symmetry:* $f(x,t) = -f(-x,t)$ for all $(x,t) \in \mathcal{M}$;
- *Equilibrium branches:* There exists a continuous function $x^\star : \mathbb{R} \to (0, d]$ with the property that $f(x,t) = 0$ in $\mathcal{M}$ if and only if $x = 0$ or $x = \pm x^\star(t)$;
- *Stability:* The origin is unstable and the equilibrium branches $\pm x^\star$ are stable, that is, for all $t \in \mathbb{R}$,

$$\begin{aligned} a(t) &:= \partial_x f(0,t) > 0 \\ a^\star(t) &:= \partial_x f(x^\star(t), t) < 0. \end{aligned} \tag{2.4}$$

- *Behaviour near $t = 0$:* We want the three equilibrium branches to come close at integer times. Given the symmetry of f, the natural assumption is that we have an "avoided pitchfork bifurcation", that is,

$$\begin{aligned} &\partial_{xxx} f(0,0) < 0 \\ &a(t) = a_0 + a_1 t^2 + \mathcal{O}(t^3), \end{aligned} \tag{2.5}$$

  where $a_1 > 0$ and $\partial_{xxx} f(0,0)$ are fixed (of order one), while $a_0 = a_0(\varepsilon) = \mathcal{O}_\varepsilon(1)$ is a positive small parameter. Is is easy to show that $x^\star(t)$ behaves like $\sqrt{a(t)}$ for small t, and admits a quadratic minimum at a time $t^\star = \mathcal{O}(a_0)$. Moreover, $a^\star(t) \asymp -a(t)$ near $t = 0$.

  We can choose a constant $T \in (0, 1/2)$ such that the derivatives of $a(t)$ and $x^\star(t)$ vanish only once in the interval $[-T, T]$. We finally require that $x^\star(t)$, $a(t)$ and $a^\star(t)$ are bounded away from zero outside this interval. We can summarize these properties as

$$x^\star(t) \asymp \begin{cases} \sqrt{a_0} & \text{for } |t| \leqslant \sqrt{a_0} \\ |t| & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \\ 1 & \text{for } T \leqslant t \leqslant 1 - T, \end{cases} \tag{2.6}$$

$$a(t) \asymp \begin{cases} a_0 & \text{for } |t| \leqslant \sqrt{a_0} \\ t^2 & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \\ 1 & \text{for } T \leqslant t \leqslant 1 - T, \end{cases} \tag{2.7}$$

$$a^\star(t) \asymp -a(t) \qquad \text{for all } t. \tag{2.8}$$

We start by considering the deterministic equation

$$\varepsilon \frac{\mathrm{d}x_t^{\det}}{\mathrm{d}t} = f(x_t^{\det}, t). \tag{2.9}$$

Without loss of generality, we may assume that x_t^{det} starts at time $-1+T$ in some $x_{-1+T}^{\mathrm{det}} > 0$. Tihonov's theorem [Gr, Ti], applied on the interval $[-1+T, -T]$, implies that x_t^{det} converges exponentially fast to a neighbourhood of order ε of $x^\star(t)$. We may thus assume that $x_{-T}^{\mathrm{det}} = x^\star(-T) + \mathcal{O}(\varepsilon)$. In fact, since $x^\star$ is decreasing at time $-T$, we may even assume that $x_{-T}^{\mathrm{det}} - x^\star(-T) \asymp \varepsilon$.

The motion of x_t^{det} in the interval $[-T, T]$ is described in the following theorem.

Theorem 2.1 (Deterministic case). *The solution x_t^{det} and the curve $x^\star(t)$ cross once and only once during the time interval $[-T, T]$. This crossing occurs at a time $\tilde{t}$ satisfying $\tilde{t} - t^\star \asymp (\varepsilon/a_0) \wedge \varepsilon^{1/3}$. There exists a constant $c_0 > 0$ such that*

$$x_t^{\mathrm{det}} - x^\star(t) \asymp \begin{cases} \dfrac{\varepsilon}{t^2} & \text{for } -T \leqslant t \leqslant -c_0(\sqrt{a_0} \vee \varepsilon^{1/3}) \\ -\dfrac{\varepsilon}{t^2} & \text{for } c_0(\sqrt{a_0} \vee \varepsilon^{1/3}) \leqslant t \leqslant T, \end{cases} \tag{2.10}$$

and thus $x_t^{\mathrm{det}} \asymp |t|$ in these time intervals. For $|t| \leqslant c_0(\sqrt{a_0} \vee \varepsilon^{1/3})$,

$$x_t^{\mathrm{det}} \asymp \begin{cases} \sqrt{a_0} & \text{if } a_0 \geqslant \varepsilon^{2/3} \\ \varepsilon^{1/3} & \text{if } a_0 \leqslant \varepsilon^{2/3}. \end{cases} \tag{2.11}$$

Finally, the linearization of f at x_t^{det} satisfies

$$\bar{a}(t) := \partial_x f(x_t^{\mathrm{det}}, t) \asymp -(t^2 \vee a_0 \vee \varepsilon^{2/3}). \tag{2.12}$$

We give the proof in Subsection 3.1. The relation (2.11) may be surprising, since it means that no matter how small we make a_0, x_t^{det} never approaches the saddle at $x = 0$ closer than a distance of order $\varepsilon^{1/3}$. This fact can be intuitively understood as follows. Even if $a_0 = 0$ and near $t = 0$, we have

$$\varepsilon \frac{\mathrm{d}x^{\mathrm{det}}}{\mathrm{d}t} \geqslant -\mathit{const}\,(x^{\mathrm{det}})^3 \qquad \Rightarrow \qquad x_t^{\mathrm{det}} \geqslant \mathit{const} \frac{x_{t_0}^{\mathrm{det}}}{\sqrt{1 + (x_{t_0}^{\mathrm{det}})^2 (t - t_0)/\varepsilon}}. \tag{2.13}$$

Since $x_{t_0}^{\mathrm{det}} \asymp \varepsilon^{1/3}$ for $t_0 \asymp -\varepsilon^{1/3}$, x_t cannot approach the origin significantly during any time interval of order $\varepsilon^{1/3}$. After such a time, however, the repulsion of the saddle will make itself felt again, preventing the solution from further approaching the origin. In other words, the time interval during which $a(t)$ is smaller than $\varepsilon^{2/3}$ is too short to allow the deterministic solution to come close to the saddle.

We return now to the SDE (2.1) with $\sigma > 0$. Assume that we start at some deterministic $x_{-1+T} > 0$. Theorem 2.3 in [BG] shows that the paths are likely to track the deterministic solution x_t^{det} with the same initial condition at a distance of order $\sigma^{1-\delta}$ for any $\delta > 0$ (with probability $\geqslant 1 - (1/\varepsilon^2)\exp\{-\mathit{const}/\sigma^{2\delta}\}$), as long as the equilibrium branches are well separated, that is, at least for $-1+T \leqslant t \leqslant -T$. A transition between the potential wells is thus unlikely if $\sigma = \mathcal{O}(|\log\varepsilon|^{-1/2\delta})$, and interesting phenomena can only be expected between the times $-T$ and T. Upon completion of one time period, i. e., at time T, the Markov property allows to repeat the above argument. Hence there is no limitation in considering the SDE (2.1) on the time interval $[-T, T]$, with a fixed initial condition x_{-T} satisfying $x_{-T} - x^\star(-T) \asymp \varepsilon$. We will denote by x_t^{det} and x_t, respectively, the solutions of (2.9) and (2.1) with the same initial condition x_{-T}.

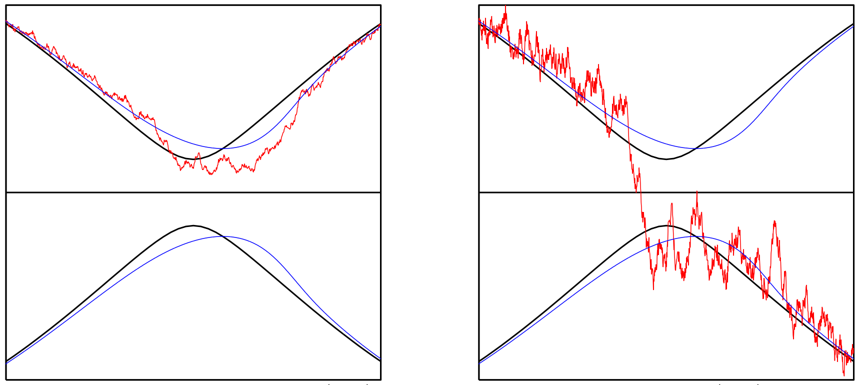

FIGURE 3. Solutions of the SDE (2.1) with symmetric drift term (2.3), shown for two different noise intensities, but for the same realization of Brownian motion. Heavy curves represent the equilibrium branches $\pm x^\star(t)$, and the straight line represents the saddle. Smooth light curves are solutions of the deterministic equation (2.9) tracking the potential wells with a small lag, while rugged curves are paths of the SDE. Parameter values are $\varepsilon = 0.01$, $a_0 = 0.02$, $\sigma = 0.02$ (left) and $\sigma = 0.08$ (right).

Let us start by describing the dynamics in a neighbourhood of $x_t^{\rm det}$. The main idea is that for σ sufficiently small, the typical spreading of paths around $x_t^{\rm det}$ should be related to the variance $v(t)$ of the solution of (2.1), linearized around $x_t^{\rm det}$. This variance is given by

$$v(t) = \frac{\sigma^2}{\varepsilon}\int_{-T}^{t} \mathrm{e}^{2\overline{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s, \qquad \text{where } \overline{\alpha}(t,s) = \int_s^t \bar{a}(u)\,\mathrm{d}u. \tag{2.14}$$

The variance is equal to zero at time $-T$, but behaves asymptotically like $\sigma^2/|2\bar{a}(t)|$. In fact, if we define the function

$$\zeta(t) := \frac{1}{2|\bar{a}(-T)|}\,\mathrm{e}^{2\overline{\alpha}(t,-T)/\varepsilon} + \frac{1}{\varepsilon}\int_{-T}^{t} \mathrm{e}^{2\overline{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s, \tag{2.15}$$

then $v(t)$ differs from $\sigma^2\zeta(t)$ by a term that becomes negligible as soon as $|\overline{\alpha}(t,-T)|$ is larger than a constant times $\varepsilon|\log\varepsilon|$. $\zeta(t)$ has the advantage to be bounded away from zero for all t, which avoids certain technical problems in the proofs. We shall show that

$$\zeta(t) \asymp \frac{1}{t^2 \vee a_0 \vee \varepsilon^{2/3}} \qquad \text{for } |t| \leqslant T. \tag{2.16}$$

We introduce the set

$$\mathcal{B}(h) = \bigl\{(x,t)\colon |t| \leqslant T, |x - x_t^{\rm det}| < h\sqrt{\zeta(t)}\bigr\}, \tag{2.17}$$

and denote by $\tau_{\mathcal{B}(h)}$ the first exit time of x_t from $\mathcal{B}(h)$.

Theorem 2.2 (Motion near the stable equilibrium branches). *There exists a constant h_0, depending only on f, such that*

- *if $-T \leqslant t \leqslant -(\sqrt{a_0} \vee \varepsilon^{1/3})$ and $h < h_0 t^2$, then*

$$\mathbb{P}^{-T,x_{-T}}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon)\exp\biggl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\biggl[1 - \mathcal{O}(\varepsilon) - \mathcal{O}\Bigl(\frac{h}{t^2}\Bigr)\biggr]\biggr\}; \tag{2.18}$$

- *if* $-(\sqrt{a_0} \vee \varepsilon^{1/3}) \leqslant t \leqslant T$ *and* $h < h_0(a_0 \vee \varepsilon^{2/3})$, *then*

$$\mathbb{P}^{-T,x_{-T}}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon)\exp\biggl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\biggl[1 - \mathcal{O}(\varepsilon) - \mathcal{O}\biggl(\frac{h}{a_0 \vee \varepsilon^{2/3}}\biggr)\biggr]\biggr\}. \tag{2.19}$$

In both cases,

$$C(t,\varepsilon) = \frac{1}{\varepsilon^2}|\overline{\alpha}(t,-T)| + 2. \tag{2.20}$$

We give the proof in Subsection 3.2. This result has several consequences. Observe first that the exponential factors in (2.18) and (2.19) are very small as soon as h is significantly larger than σ. The prefactor $C(t,\varepsilon)$ (which, unlike the exponent, we do not believe to be optimal) leads to subexponential corrections, which are negligible as soon as $h/\sigma > \mathcal{O}(|\log\varepsilon|)$. It mainly accounts for the fact that the probability for a path to leave $\mathcal{B}(h)$ increases slowly with time. The theorem shows that the typical spreading of paths around $x_t^{\det}$ is of order

$$\sigma\sqrt{\zeta(t)} \asymp \frac{\sigma}{|t| \vee \sqrt{a_0} \vee \varepsilon^{1/3}}. \tag{2.21}$$

If $\sigma \ll a_0 \vee \varepsilon^{2/3}$, we may choose $h \gg \sigma$ for all times, and thus the probability of leaving a neighbourhood of $x_t^{\det}$, let alone approach the other stable branch, is exponentially small (in $\sigma^2/(a_0 \vee \varepsilon^{2/3})^2$). On the other hand, if σ is not so small, (2.18) can still be applied to show that a transition is unlikely to occur before a time of order $-\sqrt{\sigma}$. Figure 3 illustrates this phenomenon by showing typical paths for two different noise intensities.

Let us now assume that σ is sufficiently large to allow for a transition, and examine the transition regime in more detail. We will proceed in two steps. First we will estimate the probability of *not* reaching the saddle at $x = 0$ during a time interval $[t_0, t_1]$. The symmetry of f implies that for any $t \geqslant t_1$ and $x_0 > 0$,

$$\begin{aligned}
\mathbb{P}^{t_0,x_0}\bigl\{x_t < 0\bigr\} &= \frac{1}{2}\mathbb{P}^{t_0,x_0}\bigl\{\exists s \in (t_0,t)\colon x_s = 0\bigr\} \\
&= \frac{1}{2} - \frac{1}{2}\mathbb{P}^{t_0,x_0}\bigl\{x_s > 0 \;\forall s \in [t_0,t]\bigr\} \\
&\geqslant \frac{1}{2} - \frac{1}{2}\mathbb{P}^{t_0,x_0}\bigl\{x_s > 0 \;\forall s \in [t_0,t_1]\bigr\}.
\end{aligned} \tag{2.22}$$

In the second step, we will show, independently, that paths are likely to leave a neighbourhood of $x = 0$ after time $\sqrt{\sigma}$. Thus if the probability of not reaching $x = 0$ is small, the probability of making a transition from the positive well to the negative one will be close to 1/2 (it can never exceed 1/2 because of the symmetry). This does not exclude, of course, that paths frequently switch back and forth between the two potential wells during the time interval $[-\sqrt{\sigma}, \sqrt{\sigma}]$. But it shows that (2.22) can indeed be interpreted as a lower bound on the transition probability.

Let $\delta > 0$ be a constant such that

$$x\partial_{xx}f(x,t) \leqslant 0 \qquad \text{for } |x| \leqslant \delta \text{ and } |t| \leqslant T. \tag{2.23}$$

Our hypotheses on f imply that such a δ of order one always exists. In some special cases, for instance if $f(x,t) = a(t)x - x^3$, δ may be chosen arbitrarily large.

Theorem 2.3 (Transition regime). *Let $c_1 > 0$ be a constant and assume $c_1^2\sigma \geqslant a_0 \vee \varepsilon^{2/3}$. Choose times $-T \leqslant t_0 \leqslant t_1 \leqslant T$ with $t_1 \in [-c_1\sqrt{\sigma}, c_1\sqrt{\sigma}]$, and let $h > 2\sigma$ be such that $x_s^{\det} + h\sqrt{\zeta(s)} < \delta$ for all $s \in [t_0, t_1]$. Then, if c_1 is sufficiently small and $x_0 \in (0, x_{t_0}^{\det} + \frac{1}{2}h\sqrt{\zeta(t_0)}]$,*

$$\begin{aligned}&\mathbb{P}^{t_0,x_0}\bigl\{x_s > 0\ \forall s \in [t_0,t_1]\bigr\}\\&\quad\leqslant \frac{5}{2}\Bigl(\frac{|\overline{\alpha}(t_1,t_0)|}{\varepsilon}+1\Bigr)\,\mathrm{e}^{-\bar\kappa h^2/\sigma^2} + 2\exp\Bigl\{-\bar\kappa\frac{1}{\log(h/\sigma)}\frac{\alpha(t_1,-c_1\sqrt{\sigma})}{\varepsilon}\Bigr\},\end{aligned}\tag{2.24}$$

where $\bar\kappa$ is a positive constant, and $\alpha(t,s) = \int_s^t a(u)\,\mathrm{d}u$.

The proof is given in Subsection 3.3. The first term in (2.24) is an upper bound on the probability that x_t escapes "upward". Indeed, our hypotheses on f do not exclude that other stable equilibria exist for sufficiently large x, which might trap escaping trajectories. The second term bounds the probability of x_s remaining between 0 and $x_s^{\det} + h\sqrt{\zeta(s)}$ for $-c_1\sqrt{\sigma} \leqslant s \leqslant t_1$. This estimate lies at the core of our argument, and can be understood as follows. Assume x_s starts near $x_s^{\det}$. It will perform a certain number of excursions to attempt reaching the saddle at $x = 0$. Each excursion requires a typical time of order Δs, such that $\alpha(s+\Delta s, s) \asymp \varepsilon$ (that is, $a(s)\Delta s \asymp \varepsilon$), in the sense that the probability of reaching 0 before time $s+\Delta s$ is small. After an unsuccessful excursion, x_s may exceed $x_s^{\det}$, but will return typically after another time of order Δs. Thus the total number of trials during the time interval $[-c_1\sqrt{\sigma}, t_1]$ is of order $\alpha(t_1, -c_1\sqrt{\sigma})/\varepsilon$. Under the hypotheses of the theorem, the probability of not reaching the saddle during one excursion is of order one, and thus the total number of trials determines the exponent in (2.24).

Before discussing the choice of the parameters giving an optimal bound in Theorem 2.3, let us first state the announced second step, namely the claim that the paths are likely to escape from the saddle after $t = c_1\sqrt{\sigma}$. For $\kappa \in (0,1)$, let us introduce the set

$$\mathcal{D}(\kappa) = \Bigl\{(x,t) \in [-\delta,\delta] \times [c_1\sqrt{\sigma}, T]\colon \frac{f(x,t)}{x} > \kappa a(t)\Bigr\}.\tag{2.25}$$

The upper boundary of $\mathcal{D}(\kappa)$ is a function $\tilde{x}(t) = \sqrt{1-\kappa}\,(1-\mathcal{O}(t))x^\star(t)$. Let $\tau_{\mathcal{D}(\kappa)}$ denote the first exit time of x_t from $\mathcal{D}(\kappa)$.

Theorem 2.4 (Escape from the saddle). *Let $0 < \kappa < 1$ and assume $c_1^2\sigma \geqslant a_0 \vee \varepsilon^{2/3}$. Then there exist constants $c_2 \geqslant c_1$ and $C_0 > 0$ such that*

$$\mathbb{P}^{t_2,x_2}\bigl\{\tau_{\mathcal{D}(\kappa)} \geqslant t\bigr\} \leqslant C_0\Bigl(\frac{t}{\sqrt{\sigma}}\Bigr)^2\frac{\mathrm{e}^{-\kappa\alpha(t,t_2)/2\varepsilon}}{\sqrt{1-\mathrm{e}^{-\kappa\alpha(t,t_2)/\varepsilon}}},\tag{2.26}$$

for all $(x_2,t_2) \in \mathcal{D}(\kappa)$ with $t_2 \geqslant c_2\sqrt{\sigma}$.

The proof is adapted from the proof of the similar Theorem 2.9 in [BG]. Compared to that result, we have sacrificed a factor 2 in the exponent, in order to get a weaker condition on σ. We discuss the changes in the proof in Subsection 3.4.

For the moment, let us consider $t_2 = c_2\sqrt{\sigma}$. We want to choose a t such that $\alpha(t,t_2) \geqslant \varepsilon|\log\sigma|$. Since $\alpha(t,t_2)$ is larger than a constant times $t_2^2(t-t_2)$, it suffices to choose a t of order $\sqrt{\sigma}(1+\varepsilon|\log\sigma|/\sigma^{3/2})$ for (2.26) to become small. Hence, after waiting for a time of that order, we find

$$\mathbb{P}^{t_2,x_2}\bigl\{\tau_{\mathcal{D}(\kappa)} \geqslant t\bigr\} \leqslant \mathit{const}\,|\log\sigma|\sigma^{\kappa/2},\tag{2.27}$$

which shows that most trajectories will have left $\mathcal{D}(\kappa)$ by time t, see Fig. 4.

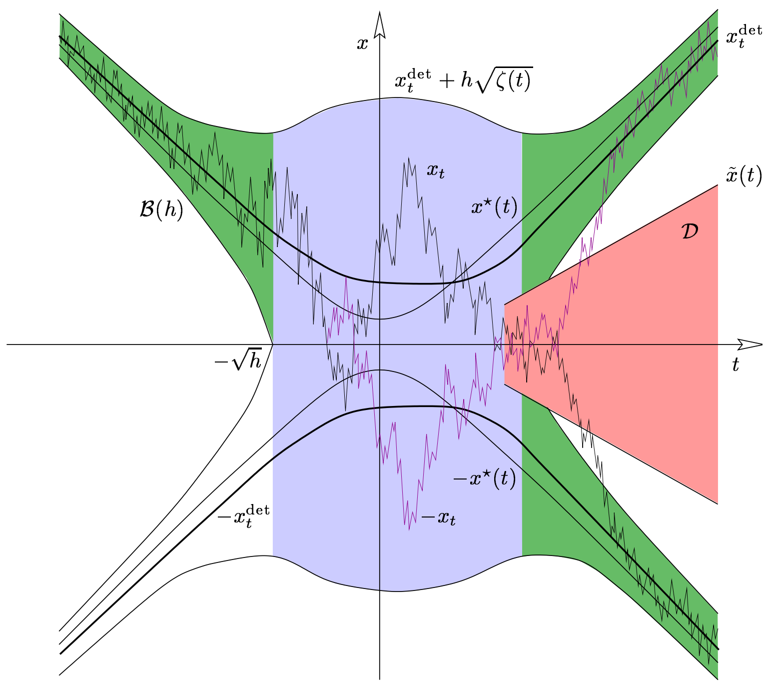

FIGURE 4. A typical path x_t of the SDE (2.1) in the symmetric case, shown in a neighbourhood of the origin, where the potential barrier reaches its minimal height. We show here a situation where the noise intensity is large enough to allow a transition. The potential wells at $\pm x^\star(t)$ behave like $\pm(\sqrt{a_0} \vee |t|)$. The deterministic solution $x_t^{\rm det}$ starting near the right-hand potential well tracks $x^\star(t)$ at a distance at most of order $(\varepsilon/a_0) \vee \varepsilon^{1/3}$, and never approaches the saddle at $x = 0$ closer than $\sqrt{a_0} \vee \varepsilon^{1/3}$. The path x_t is likely to stay in the set $\mathcal{B}(h)$ up to time $-\sqrt{h}$ when $h \gg \sigma$. Between times $-\sqrt{\sigma}$ and $\sqrt{\sigma}$, the path is likely to reach the origin. It may continue to jump back and forth between the potential wells up to time $\sqrt{\sigma}$, but is likely to leave a neighbourhood $\mathcal{D}$ of the saddle for times slightly larger than $\sqrt{\sigma}$. For each realization ω such that $x_t(\omega)$ reaches the saddle at a time τ, there is a realization ω' such that $x_t(\omega') = -x_t(\omega)$ is the mirror image of x_t for $t \geqslant \tau$, which explains why the probability to choose one well or the other after the transition region is close to $1/2$.

It remains to show that the paths are likely to approach either $x^\star(t)$ or $-x^\star(t)$ after leaving $\mathcal{D}(\kappa)$. Let us first consider the solution $\widehat{x}_t^{\rm det}$ of the deterministic differential equation (2.9) with initial time $t_2 \geqslant c_2\sqrt{\sigma}$ and initial condition $|x_2 - x^\star(t)| \leqslant ct_2$ for some small constant $c > 0$. Here we need to choose c small in order to arrange for $\widehat{a}(t) = \partial_x f(\widehat{x}_t^{\rm det}, t) \asymp -t^2$ which allows us to proceed as in our investigation of the motion for $t \leqslant -(\sqrt{a}_0 \vee \varepsilon^{1/3})$, cf. Subsections 3.1 and 3.2. Under these assumptions, $\widehat{x}_t^{\rm det}$ approaches a neighbourhood of $\pm x^\star(t)$ exponentially fast and then tracks the equilibrium branch at distance ε/t^2. As before, one can show that the path x_t of the solution of the SDE (2.1) with the same initial condition is likely to remain in a strip around $\widehat{x}_t^{\rm det}$ of width scaling with $\sigma/\sqrt{\widehat{a}(t)} \asymp \sigma/t$. So if a path x_t leaves $\mathcal{D}(\kappa)$ at time $\tau_{\mathcal{D}(\kappa)}$, then this path is likely to approach $x^\star(t)$, if $x_{\tau_{\mathcal{D}(\kappa)}}$ is positive, and $-x^\star(t)$, otherwise. Note that $|x^\star(\tau_{\mathcal{D}(\kappa)}) - x_{\tau_{\mathcal{D}(\kappa)}}|$ has to be smaller than $c\tau_{\mathcal{D}(\kappa)}$, which restricts the possible values for κ. Therefore, we choose κ small enough to guarantee that $x^\star(t) - \tilde{x}(t) \leqslant ct$ for all $t \geqslant c_1\sqrt{\sigma}$. Finally note that paths which are not in $\mathcal{D}(\kappa)$ at time $c_2\sqrt{\sigma}$ but are not further away from $\pm x^\star$ than ct at some time t will also approach the corresponding equilibrium branch.

Let us now discuss the choice of the parameters in (2.24) giving an optimal bound. For $\sigma \geqslant (a_0 \vee \varepsilon^{2/3})/c_1^2$, Theorem 2.2 shows that up to time t slightly less than $-\sqrt{\sigma}$, the paths are concentrated around $x_t^{\det}$. Therefore, here we should choose an initial time t_0 slightly before $-\sqrt{\sigma}$. In order for the second term in (2.24) to be small, we want to choose $t_1 - (-c_1\sqrt{\sigma})$ as large as possible. Note, however, that $h\sqrt{\zeta(s)}$ has to be smaller than $\delta - x_s^{\det}$ for all s. Since the order of $\zeta(s)$ is increasing for $s \leqslant -(\sqrt{a_0} \vee \varepsilon^{1/3})$, it turns out to be more advantageous to choose t_1 negative and of order $-\sqrt{\sigma}$, say $t_1 = -\frac{1}{2}c_1\sqrt{\sigma}$. In this case, $|\alpha(t_1, -c_1\sqrt{\sigma})|$ is larger than a constant times $\sigma^{3/2}$ (independently of a_0), and this does not improve significantly for larger admissible t_1. At the same time, this choice allows us to take $h \asymp \delta\sqrt{\sigma}$. We find

$$\begin{aligned}\mathbb{P}^{t_0,x_0}\Big\{x_s > 0 \;\forall s \in [t_0, c_1\sqrt{\sigma}\,]\Big\} &\leqslant \mathbb{P}^{t_0,x_0}\Big\{x_s > 0 \;\forall s \in [t_0, -\tfrac{1}{2}c_1\sqrt{\sigma}\,]\Big\} \\ &\leqslant \left(\frac{|t_0|}{\sqrt{\sigma}}\right)^3 \frac{\sigma^{3/2}}{\varepsilon}\,\mathrm{e}^{-\mathcal{O}(\delta^2/\sigma)} + \exp\Big\{-\frac{const}{\log(\delta^2/\sigma)}\frac{\sigma^{3/2}}{\varepsilon}\Big\}.\end{aligned} \tag{2.28}$$

Consider first the generic case $\delta \asymp 1$. The second term in (2.28) becomes small as soon as $\sigma/(|\log\sigma|)^{2/3} \gg \varepsilon^{2/3}$ holds in addition to the general condition $c_1^2\sigma \geqslant a_0 \vee \varepsilon^{2/3}$. The first term is small as long as $\sigma = \mathcal{O}(1/\log(\sigma/\varepsilon^{2/3}))$. We thus obtain the following regimes:

- for $\sigma \leqslant a_0 \vee \varepsilon^{2/3}$, the transition probability is exponentially small in $\sigma^2/(a_0 \vee \varepsilon^{2/3})^2$;
- for $\sigma \geqslant a_0/c_1^2$ with $(\varepsilon|\log\varepsilon|)^{2/3} \ll \sigma \ll 1/|\log\varepsilon|$, the probability of a transition between the wells is exponentially close to $1/2$, with an exponent given essentially (up to logarithmic corrections) by

$$\frac{\sigma^{3/2}}{\varepsilon} \wedge \frac{1}{\sigma}; \tag{2.29}$$

- for $\sigma \geqslant 1/|\log\varepsilon|$, the paths become so poorly localized that it is no longer meaningful to speak of a transition probability.

(2.29) shows that the transition probability becomes optimal for $\sigma \asymp \varepsilon^{2/5}$. For larger values of the noise intensity, the possibility of paths escaping "upward" becomes sufficiently important to decrease the transition probability. However, if the function f is such that δ can be chosen arbitrarily large, the second term in (2.29) can be removed without changing the first one (up to logarithmic corrections) by taking $\delta^2 = \sigma/\varepsilon$, for instance. In that case, transitions between the wells become the more likely the larger the ratio $\sigma/\varepsilon^{2/3}$ is.

One should note that a typical path will reach maximal values of the order $\sigma\sqrt{\zeta(0)} \asymp \sigma/(\sqrt{a_0} \vee \varepsilon^{1/3})$. Thus, due to the flatness of the potential near $t = 0$, if σ is larger than $\sqrt{a_0} \vee \varepsilon^{1/3}$, the spreading of the paths during the transition interval is larger than the maximal distance between the wells away from the transition. In general we cannot exclude that paths escape to other attractors, if the potential has more than two wells.

It may be surprising that the order of the transition probability is independent of a_0 as soon as $\sigma > a_0$. Intuitively, one would rather expect this probability to depend on the ratio a_0^2/σ^2, because of Kramers' law. The fact that this is not the case illustrates the necessity of a good understanding of dynamical effects (as opposed to a quasistatic picture). Although the potential barrier is smallest between the times $-\sqrt{a_0}$ and $\sqrt{a_0}$, the paths have more opportunities to reach the saddle during larger time intervals. The optimal time interval turns out to have a length of the order $\sqrt{\sigma}$, which corresponds to the regime where diffusive behaviour prevails over the influence of the drift.

2.3 Asymmetric case

We consider in this subsection the SDE (2.1) in the case of f being periodic in t and admitting two stable equilibrium branches, but without the symmetry assumption. Instead, we want each of the potential wells to approach the saddle once in every time period, but at different times for the left-hand and the right-hand potential well. A typical example of such a function is

$$f(x,t) = x - x^3 + \lambda(t) \qquad \text{with} \qquad \lambda(t) = -(\lambda_c - a_0)\cos 2\pi t. \tag{2.30}$$

Here $\lambda_c = 2/(3\sqrt{3})$ is defined by the fact that f has two stable equilibria if and only if $|\lambda| < \lambda_c$. Observe that $\partial_x f$ vanishes at $x = \pm x_c = \pm 1/\sqrt{3}$, and

$$\begin{aligned} f(x_c + y, t) &= \lambda_c - (\lambda_c - a_0)\cos 2\pi t - \sqrt{3}y^2 - y^3 \\ &= a_0 + 2\pi^2(\lambda_c - a_0)t^2 + \mathcal{O}(t^4) - \sqrt{3}y^2 - y^3. \end{aligned} \tag{2.31}$$

Here the function $f(x_c, t)$ plays the role that $a(t)$ played in the symmetric case, and near $t = 0$ the right-hand potential well and the saddle behave like $x_c \pm 3^{-1/4}\sqrt{f(x_c,t)}$, while the left-hand potential well is isolated. Near $t = 1/2$, a similar close encounter takes place between the saddle and the left-hand potential well.

We will consider a more general class of functions $f : \mathbb{R}^2 \to \mathbb{R}$, which we assume to satisfy the following hypotheses:

- *Smoothness:* $f \in \mathcal{C}^3(\mathcal{M}, \mathbb{R})$, where $\mathcal{M} = [-d,d] \times \mathbb{R}$ and $d > 0$ is a constant;
- *Periodicity:* $f(x, t+1) = f(x,t)$ for all $(x,t) \in \mathcal{M}$;
- *Equilibrium branches:* There exist continuous functions $x^\star_- < x^\star_0 < x^\star_+$ from $\mathbb{R}$ to $[-d,d]$ with the property that $f(x,t) = 0$ in $\mathcal{M}$ if and only if $x = x^\star_\pm(t)$ or $x = x^\star_0(t)$; the zeroes of f should be isolated in the following sense: for every $\delta > 0$, there should exist a constant $\rho > 0$ such that, if $|x - x^\star_\pm(t)| \geqslant \delta$ and $|x - x^\star_0(t)| \geqslant \delta$, then $|f(x,t)| \geqslant \rho$.[2]
- *Stability:* The equilibrium branches $x^\star_\pm$ are stable and the equilibrium branch $x^\star_0$ is unstable, that is, for all $t \in \mathbb{R}$,

$$\begin{aligned} a^\star_\pm(t) &:= \partial_x f(x^\star_\pm(t), t) < 0 \\ a^\star_0(t) &:= \partial_x f(x^\star_0(t), t) > 0. \end{aligned} \tag{2.32}$$

- *Behaviour near $t = 0$:* We want $x^\star_+$ and $x^\star_0$ to come close at integer times. Here the natural assumption is that we have an "avoided saddle–node bifurcation", that is, there exists an $x_c \in (-\delta, \delta)$ such that

$$\begin{aligned} \partial_{xx} f(x_c, 0) &< 0 \\ \partial_x f(x_c, t) &= \mathcal{O}(t^2) \\ f(x_c, t) &= a_0 + a_1 t^2 + \mathcal{O}(t^3), \end{aligned} \tag{2.33}$$

  where $a_1 > 0$ and $\partial_{xx} f(x_c, 0)$ are fixed (of order one), while $a_0 = a_0(\varepsilon) = o_\varepsilon(1)$ is a positive small parameter. These assumptions imply that $x^\star_+(t)$ reaches a local minimum at a time $t^\star_+ = \mathcal{O}(a_0)$, and $x^\star_0(t)$ reaches a local maximum at a possibly

[2] Since f depends on a small parameter a_0, we want to avoid that $f(x,t)$ approaches zero elsewhere but near the three equilibrium branches, even when a_0 becomes small.

different time $t_0^\star = \mathcal{O}(a_0)$. We can assume that for a sufficiently small constant $T > 0$, the three equilibrium branches and the linearization of f around them satisfy

$$
\begin{aligned}
x_+^\star(t) - x_c &\asymp \begin{cases} \sqrt{a_0} & \text{for } |t| \leqslant \sqrt{a_0} \\ |t| & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \end{cases} & a_+^\star(t) &\asymp \begin{cases} -\sqrt{a_0} & \text{for } |t| \leqslant \sqrt{a_0} \\ -|t| & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \end{cases} \\
x_0^\star(t) - x_c &\asymp \begin{cases} -\sqrt{a_0} & \text{for } |t| \leqslant \sqrt{a_0} \\ -|t| & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \end{cases} & a_0^\star(t) &\asymp \begin{cases} \sqrt{a_0} & \text{for } |t| \leqslant \sqrt{a_0} \\ |t| & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \end{cases} \\
x_-^\star(t) - x_c &\asymp -1 \qquad \text{for } |t| \leqslant T & a_-^\star(t) &\asymp -1 \qquad \text{for } |t| \leqslant T.
\end{aligned} \tag{2.34}
$$

- *Behaviour near $t = t_c$:* We want $x_-^\star$ and $x_0^\star$ to come close at some time $t_c \in (T, 1-T)$. This is achieved by assuming that similar relations as (2.33), but with opposite signs, hold at a point (x_c', t_c).
- *Behaviour between the close encounters:* To exclude the possibility of other almost-bifurcations, we require that $x_+^\star(t) - x_0^\star(t)$ and $x_0^\star(t) - x_-^\star(t)$, as well as the derivatives (2.32), are bounded away from zero for $T < t < t_c - T$ and $t_c + T < t < 1 - T$.

Note that a sufficient assumption for the requirements on the behaviour near (x_c', t_c) to hold is that $f(x, t + \frac{1}{2}) = -f(-x, t)$ for all (x, t).

We start by considering the deterministic equation

$$
\varepsilon \frac{\mathrm{d}x_t^{\det}}{\mathrm{d}t} = f(x_t^{\det}, t). \tag{2.35}
$$

As in the symmetric case, it is sufficient to consider the dynamics in the time interval $[-T, T]$, with an initial condition satisfying $x_{-T}^{\det} - x_+^\star(-T) \asymp \varepsilon$. The situation in the time interval $[t_c - T, t_c + T]$ can be described in exactly the same way.

Theorem 2.5 (Deterministic case). *The solution $x_t^{\det}$ and the curve $x_+^\star(t)$ cross once and only once during the time interval $[-T, T]$. This crossing occurs at a time $\tilde{t}$ satisfying $\tilde{t} - t_+^\star \asymp (\varepsilon/\sqrt{a_0}) \wedge \sqrt{\varepsilon}$. There exists a constant $c_0 > 0$ such that*

$$
x_t^{\det} - x_+^\star(t) \asymp \begin{cases} \dfrac{\varepsilon}{|t|} & \text{for } -T \leqslant t \leqslant -c_0(\sqrt{a_0} \vee \sqrt{\varepsilon}) \\ -\dfrac{\varepsilon}{|t|} & \text{for } c_0(\sqrt{a_0} \vee \sqrt{\varepsilon}) \leqslant t \leqslant T, \end{cases} \tag{2.36}
$$

and thus $x_t^{\det} - x_c \asymp |t|$ in these time intervals. For $|t| \leqslant c_0(\sqrt{a_0} \vee \sqrt{\varepsilon})$,

$$
x_t^{\det} - x_c \asymp \begin{cases} \sqrt{a_0} & \text{if } a_0 \geqslant \varepsilon \\ \sqrt{\varepsilon} & \text{if } a_0 \leqslant \varepsilon. \end{cases} \tag{2.37}
$$

The linearization of f at $x_t^{\det}$ satisfies

$$
\bar{a}(t) := \partial_x f(x_t^{\det}, t) \asymp -(|t| \vee \sqrt{a_0} \vee \sqrt{\varepsilon}). \tag{2.38}
$$

Moreover, (2.35) *admits a particular solution $\widehat{x}_t^{\det}$ tracking the* unstable *equilibrium branch $x_0^\star(t)$. It satisfies analogous relations, namely, $\widehat{x}_t^{\det}$ and $x_0^\star(t)$ cross once at a time $\hat{t}$ satisfying $\hat{t} - t_0^\star \asymp -(\tilde{t} - t_+^\star)$, and* (2.36), (2.37) *and* (2.38) *hold for $\widehat{x}^{\det}$ and $x_0^\star(t)$, but with opposite signs.*

The proof is similar to the proof of Theorem 2.1, and we comment on a few minor differences in Subsection 4.1. Note that (2.37) implies that $x_t^{\det}$ never approaches the saddle at $x_0^\star(t)$ closer than a distance of order $\sqrt{\varepsilon}$.

We return now to the SDE (2.1) with $\sigma > 0$. We will denote by $x_t^{\det}$ and x_t, respectively, the solutions of (2.35) and (2.1) with the same initial condition x_{-T} satisfying $x_{-T} - x_+^\star(-T) \asymp \varepsilon$. We introduce again the function

$$\zeta(t) := \frac{1}{2|\bar{a}(-T)|}\,\mathrm{e}^{2\overline{\alpha}(t,-T)/\varepsilon} + \frac{1}{\varepsilon}\int_{-T}^{t} \mathrm{e}^{2\overline{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s, \qquad \text{where } \overline{\alpha}(t,s) = \int_s^t \bar{a}(u)\,\mathrm{d}u, \tag{2.39}$$

which behaves, in this case, like

$$\zeta(t) \asymp \frac{1}{|t| \vee \sqrt{a_0} \vee \sqrt{\varepsilon}} \qquad \text{for } |t| \leqslant T. \tag{2.40}$$

We define once more the set

$$\mathcal{B}(h) = \bigl\{(x,t) \colon |t| \leqslant T, |x - x_t^{\det}| < h\sqrt{\zeta(t)}\bigr\}, \tag{2.41}$$

and denote by $\tau_{\mathcal{B}(h)}$ the first exit time of x_t from $\mathcal{B}(h)$.

Theorem 2.6 (Motion near the stable equilibrium branches). *There exists a constant h_0, depending only on f, such that*

- *if $-T \leqslant t \leqslant -(\sqrt{a_0} \vee \sqrt{\varepsilon})$ and $h < h_0|t|^{3/2}$, then*

$$\mathbb{P}^{-T,x_{-T}}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon)\exp\biggl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\biggl[1 - \mathcal{O}(\varepsilon) - \mathcal{O}\biggl(\frac{h}{t^{3/2}}\biggr)\biggr]\biggr\}; \tag{2.42}$$

- *if $-(\sqrt{a_0} \vee \sqrt{\varepsilon}) \leqslant t \leqslant T$ and $h < h_0(a_0^{3/4} \vee \varepsilon^{3/4})$, then*

$$\mathbb{P}^{-T,x_{-T}}\bigl\{\tau_{\mathcal{B}(h)} < t\bigr\} \leqslant C(t,\varepsilon)\exp\biggl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\biggl[1 - \mathcal{O}(\varepsilon) - \mathcal{O}\biggl(\frac{h}{a_0^{3/4} \vee \varepsilon^{3/4}}\biggr)\biggr]\biggr\}. \tag{2.43}$$

In both cases,

$$C(t,\varepsilon) = \frac{1}{\varepsilon^2}|\overline{\alpha}(t,-T)| + 2. \tag{2.44}$$

This result is proved in exactly the same way as Theorem 2.2. It has similar consequences, only with different values of the exponents. The typical spreading of paths around $x_t^{\det}$ is of order

$$\sigma\sqrt{\zeta(t)} \asymp \frac{\sigma}{\sqrt{|t|} \vee a_0^{1/4} \vee \varepsilon^{1/4}}. \tag{2.45}$$

If $\sigma \ll a_0^{3/4} \vee \varepsilon^{3/4}$, the probability of leaving a neighbourhood of $x_t^{\det}$, or making a transition to the other stable equilibrium branch, is exponentially small (in $\sigma^2/(a_0^{3/2} \vee \varepsilon^{3/2})$). On the other hand, if σ is not so small, (2.42) can still be applied to show that a transition is unlikely to occur before a time of order $-\sigma^{2/3}$.

Let us now assume that σ is sufficiently large for a transition to take place, i.e. that $\sigma \geqslant a_0^{3/4} \vee \varepsilon^{3/4}$. We want to give an upper bound on the probability *not* to make a transition. Let us introduce levels $\delta_0 < \delta_1 < x_{\mathrm{c}} < \delta_2$ such that

$$\begin{aligned} f(x,t) &\asymp -1 && \text{for } \delta_0 \leqslant x \leqslant \delta_1 \text{ and } |t| \leqslant T \\ \partial_{xx} f(x,t) &\leqslant 0 && \text{for } \delta_1 \leqslant x \leqslant \delta_2 \text{ and } |t| \leqslant T. \end{aligned} \tag{2.46}$$

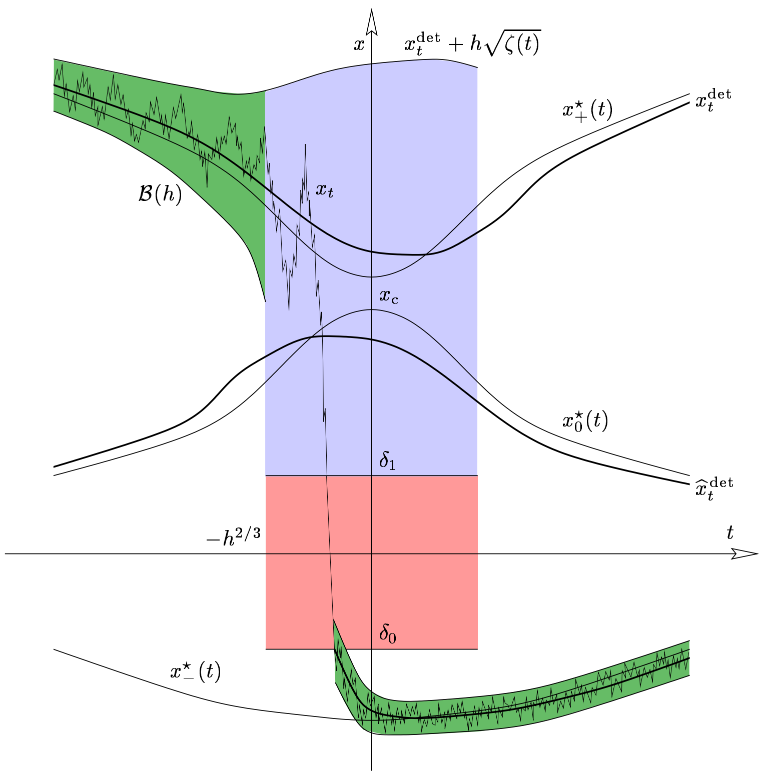

FIGURE 5. A typical path x_t of the SDE (2.1) in the asymmetric case, shown near $t = 0$, where the right-hand well at $x_+^\star$ approaches the saddle at $x_0^\star$. We show again a situation where the noise intensity is large enough to allow a transition. The deterministic solution $x_t^{\rm det}$ starting near the right-hand potential well tracks $x_+^\star(t)$ at a distance at most of order $(\varepsilon/\sqrt{a_0}) \wedge \sqrt{\varepsilon}$, and never approaches the saddle at $x_0^\star$ closer than $\sqrt{a_0} \vee \sqrt{\varepsilon}$. The path x_t is likely to stay in the set $\mathcal{B}(h)$ up to time $-h^{2/3}$ when $h \gg \sigma$. Between times $-h^{2/3}$ and $h^{2/3}$, the path is likely to reach the saddle. Although it may fall back into the right-hand potential well, it is likely to finally overcome the potential barrier and reach a level δ_1 of order 1 below the saddle, after which it quickly reaches a lower level δ_0. The distance between δ_1 and δ_0 can be much larger than in this picture. Finally, the path will track the deterministic solution starting in $x = \delta_0$, which approaches a neighbourhood of order ε of the left-hand potential well at $x_-^\star$.

Here δ_0 and δ_1 are always of order 1 (in fact, we must have $\delta_0 > x_-^\star(t)$ for all $t \in [-T, T]$), and we think of δ_0 as being in the basin of attraction of $x_-^\star$. Our hypotheses imply that a δ_2 of order one satisfying (2.46) always exists, but δ_2 may be chosen arbitrarily large in particular cases such as $f(x,t) = x - x^3 + \lambda(t)$.

The non-transition probability can be estimated by distinguishing three cases:

- Either x_t, starting in $x_0 > x_{\rm c}$ at some $t_0 < 0$, never reaches δ_1. The probability of this event can be shown to be small in a similar way as in Theorem 2.3, the main difference being that due to the asymmetry, we can do better than estimating the probability not to reach the saddle.
- If x_t reaches δ_1, one can estimate in a very simple way the probability not to reach δ_0 as well, using the fact that the drift term is bounded away from zero.
- If x_t reaches δ_0, Theorem 2.3 in [BG] shows that x_t is likely to reach a small neighbourhood of $x_-^\star$ as well.

Theorem 2.7 (Transition regime). *Let c_1 and c_2 be positive constants and assume that $c_1^{3/2}\sigma \geqslant a_0^{3/4} \vee \varepsilon^{3/4}$. Choose times $-T \leqslant t_0 \leqslant t_1 \leqslant t \leqslant T$ with $t_1 \in [-c_1\sigma^{2/3}, c_1\sigma^{2/3}]$ and $t \geqslant t_1 + c_2\varepsilon$. Let $h > 2\sigma$ be such that $x_s^{\det} + h\sqrt{\zeta(s)} < \delta_2$ for all $s \in [t_0, t_1]$. Then, for sufficiently small c_1, sufficiently large c_2 and all $x_0 \in (\delta_1, x_{t_0}^{\det} + \frac12 h\sqrt{\zeta(t_0)}]$,*

$$\begin{aligned}\mathbb{P}^{t_0,x_0}\bigl\{x_s > \delta_0\ \forall s \in [t_0,t]\bigr\} \leqslant{}& \frac52\Bigl(\frac{|\overline{\alpha}(t_1,t_0)|}{\varepsilon} + 1\Bigr)\,\mathrm{e}^{-\kappa h^2/\sigma^2} \\ &+ \frac32\exp\Bigl\{-\kappa\frac{1}{\log(h/\sigma)\vee|\log\sigma|}\frac{\widehat{\alpha}(t_1,-c_1\sigma^{2/3})}{\varepsilon}\Bigr\} \\ &+ \mathrm{e}^{-\kappa/\sigma^2},\end{aligned} \tag{2.47}$$

where κ is a positive constant, and $\widehat{\alpha}(t,s) = \int_s^t \partial_x f(\widehat{x}_u^{\det}, u)\,\mathrm{d}u$.

The proof is given in Subsection 4.2. The three terms on the right-hand side of (2.47) bound, respectively, the probability that x_t escapes through the upper boundary $x_s^{\det} + h\sqrt{\zeta(s)}$, the probability that x_t reaches neither the upper boundary nor δ_1, and the probability that x_t does not reach δ_0 when starting on δ_1 (Fig. 5). The crucial term is the second one.

Let us now discuss the optimal choice of parameters. If we choose $t_1 = -\frac12 c_1\sigma^{2/3}$, we can take $h \asymp \tilde\delta_2\sigma^{1/3}$, where $\tilde\delta_2 = \delta_2 - x_{\mathrm{c}}$, and we get the estimate

$$\begin{aligned}&\mathbb{P}^{t_0,x_0}\Bigl\{x_s > \delta_0\ \forall s \in [t_0,t]\Bigr\} \\ &\qquad\leqslant \frac{t_0^2}{\varepsilon}\,\mathrm{e}^{-\mathcal{O}(\tilde\delta_2^2/\sigma^{4/3})} + \exp\Bigl\{-\frac{const}{\log(\tilde\delta_2^2/\sigma^{4/3})\vee|\log\sigma|}\frac{\sigma^{4/3}}{\varepsilon}\Bigr\} + \mathrm{e}^{-\kappa/\sigma^2}\,.\end{aligned} \tag{2.48}$$

As in the symmetric case, when $\delta_2 \asymp 1$, we obtain the following regimes:

- for $\sigma \leqslant a_0^{3/4}\vee\varepsilon^{3/4}$, the transition probability is exponentially small in $\sigma^2/(a_0^{3/4}\vee\varepsilon^{3/4})^2$;
- for $a_0^{3/4} \vee \varepsilon^{3/4} \ll \sigma \ll (1/|\log\varepsilon|)^{3/4}$, the transition probability is exponentially close to 1, with an exponent given essentially (up to logarithmic corrections) by

$$\frac{\sigma^{4/3}}{\varepsilon} \wedge \frac{1}{\sigma^{4/3}}; \tag{2.49}$$

- for $\sigma \geqslant 1/|\log\varepsilon|$, the paths become so poorly localized that it is no longer meaningful to speak of a transition probability.

The transition probability becomes optimal for $\sigma \asymp \varepsilon^{3/8}$. Note, once again, that the exponent is independent of a_0.

If the function f is such that δ_2 can be chosen arbitrarily large, the second term in (2.49) can be removed without changing the first one (up to logarithmic corrections) by taking $\tilde\delta_2^2 = \sigma^{8/3}/\varepsilon$ for instance. If $\sigma > a_0^{1/4}\vee\varepsilon^{1/4}$, the paths may become extremely delocalised in the transition zone, and could escape to other attractors.

3 Symmetric case

We consider in this section the nonlinear SDE

$$dx_t = \frac{1}{\varepsilon} f(x_t, t) \, dt + \frac{\sigma}{\sqrt{\varepsilon}} \, dW_t, \tag{3.1}$$

where f satisfies the hypotheses given at the beginning of Subsection 2.2. By rescaling x, we can arrange for $\partial_{xxx} f(0,0) = -6$, so that Taylor's formula allows us to write

$$\begin{aligned} f(x,t) &= x\big[a(t) + g_0(x,t)\big] \\ \partial_x f(x,t) &= a(t) + g_1(x,t), \end{aligned} \tag{3.2}$$

where $g_0, g_1 \in \mathcal{C}^3$ satisfy

$$\begin{aligned} g_0(x,t) &= \big[-1 + r_0(x,t)\big] x^2 \\ g_1(x,t) &= \big[-3 + r_1(x,t)\big] x^2, \end{aligned} \tag{3.3}$$

with continuously differentiable functions r_0, r_1 satisfying $r_0(0,0) = r_1(0,0) = 0$.

The implicit function theorem shows the existence, for small t, of an equilibrium curve

$$x^\star(t) = \big[1 + \mathcal{O}\big(\sqrt{a(t)}\,\big)\big] \sqrt{a(t)}. \tag{3.4}$$

For small t, the curve $x^\star(t)$ behaves like $\sqrt{a_0}\big[1 + \mathcal{O}((t/\sqrt{a_0})^2) + \mathcal{O}(\sqrt{a_0})\big]$, and it admits a quadratic minimum at some time $t^\star = \mathcal{O}(a_0)$. Thus we can choose a constant $T \in (0, 1/2)$ such that

$$x^\star(t) \asymp \begin{cases} \sqrt{a_0} & \text{for } |t| \leqslant \sqrt{a_0} \\ |t| & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \\ 1 & \text{for } T \leqslant t \leqslant 1 - T. \end{cases} \tag{3.5}$$

3.1 Deterministic case

In this subsection we consider the deterministic equation

$$\varepsilon \frac{dx_t}{dt} = f(x_t, t). \tag{3.6}$$

As already mentioned, Tihonov's theorem allows us to restrict the analysis to the time interval $[-T, T]$, and to assume that $x_{-T} - x^\star(-T) \asymp \varepsilon$.

Remark 3.1. During the time interval $[-T, T]$, the process x_t crosses the equilibrium branch $x^\star(t)$ once and only once, the time $\tilde{t}$ of the crossing satisfying $\tilde{t} \geqslant t^\star$. This fact is due to the property that x_t is strictly decreasing when lying above $x^\star(t)$, and strictly increasing when lying below. Let $\tilde{t}_1 < \tilde{t}_2 < \dots$ be the times of the successive crossings of x_t and $x^\star(t)$ in $[-T, T]$. Then x_t is decreasing between $-T$ and $\tilde{t}_1$ (since $x_{-T} - x^\star(-T) > 0$), increasing for $\tilde{t}_1 < t < \tilde{t}_2$, and so on. Thus $x^\star(t)$ must be increasing for t slightly larger than $\tilde{t}_1$, and decreasing for t slightly larger than $\tilde{t}_2$. Since, by assumption, $x^\star(t)$ is decreasing on $[-T, t^\star)$ and increasing on $(t^\star, T]$, this implies that $\tilde{t}_1 \geqslant t^\star$ and $\tilde{t}_2 > T$. Therefore, there is at most one crossing. We shall see below that x_t and $x^\star(t)$ actually cross and we will also determine the order of that time $\tilde{t}$.

We consider now the difference $y_t = x_t - x^\star(t)$. It satisfies the equation

$$\varepsilon \frac{\mathrm{d}y}{\mathrm{d}t} = a^\star(t)y + b^\star(y,t) - \varepsilon \frac{\mathrm{d}x^\star}{\mathrm{d}t}, \tag{3.7}$$

where Taylor's formula, (3.2) and (3.3) yield the relations

$$a^\star(t) = -2a(t)\bigl[1 + \mathcal{O}\bigl(\sqrt{a(t)}\,\bigr)\bigr] \asymp \begin{cases} -a_0 & \text{for } |t| \leqslant \sqrt{a_0} \\ -t^2 & \text{for } \sqrt{a_0} \leqslant |t| \leqslant T \end{cases} \tag{3.8}$$

$$b^\star(y,t) = -\bigl(3x^\star(t) + y\bigr)y^2\bigl[1 + \mathcal{O}(x^\star(t) + y)\bigr] \tag{3.9}$$

$$\frac{\mathrm{d}x^\star}{\mathrm{d}t}(t) \asymp \begin{cases} -1 & \text{for } -T \leqslant t \leqslant -\sqrt{a_0} \\ \dfrac{(t-t^\star)}{\sqrt{a_0}} & \text{for } |t| \leqslant \sqrt{a_0} \\ 1 & \text{for } \sqrt{a_0} \leqslant t \leqslant T, \end{cases} \tag{3.10}$$

with $t^\star = t^\star(a_0) = \mathcal{O}(a_0)$.

We start by giving a technical result that we will need several times.

Lemma 3.2. *Let $\tilde{a}(t)$ be a continuous function satisfying $\tilde{a}(t) \asymp -(\beta \vee t^2)$ for $|t| \leqslant T$, where $\beta = \beta(\varepsilon) \geqslant 0$. Let $\chi_0 \asymp 1$, and define $\tilde{\alpha}(t,s) = \int_s^t \tilde{a}(u)\,\mathrm{d}u$. Then*

$$\chi_0\,\mathrm{e}^{\tilde{\alpha}(t,-T)/\varepsilon} + \frac{1}{\varepsilon}\int_{-T}^t \mathrm{e}^{\tilde{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s \asymp \begin{cases} \dfrac{1}{\beta \vee \varepsilon^{2/3}} & \textit{for } |t| \leqslant \sqrt{\beta} \vee \varepsilon^{1/3} \\ \dfrac{1}{t^2} & \textit{for } \sqrt{\beta} \vee \varepsilon^{1/3} \leqslant |t| \leqslant T. \end{cases} \tag{3.11}$$

Proof: To prove the lemma, we take advantage of the fact that the expression on the left-hand side of (3.11) is the solution of an ordinary differential equation. By the semi-group property, we may consider separately the following regimes: For $\beta \geqslant \varepsilon^{2/3}$, we distinguish the cases $t \in [-T,-T/2]$, $[-T/2,-\sqrt{\beta}]$, $[-\sqrt{\beta},-\varepsilon/a_0]$, $[-\varepsilon/a_0,\varepsilon/a_0]$, $[\varepsilon/a_0,\sqrt{\beta}]$, $[\sqrt{\beta},T]$ and for $\beta < \varepsilon^{2/3}$, we deal separately with $t \in [-T,-T/2]$, $[-T/2,-\varepsilon^{1/3}]$, $[-\varepsilon^{1/3},-\sqrt{\beta}]$, $[-\sqrt{\beta},\sqrt{\beta}]$, $[\sqrt{\beta},\varepsilon^{1/3}]$, $[\varepsilon^{1/3},T]$. On each of these time intervals the claimed behaviour follows easily by elementary calculus, see also the similar result [BG, Lemma 4.2]. □

Proposition 3.3. *There exists a constant $c_0 > 0$ such that the solution of (3.7) with initial condition $y_{-T} \asymp \varepsilon$ satisfies*

$$y_t \asymp \frac{\varepsilon}{t^2} \qquad \textit{for } -T \leqslant t \leqslant -c_0(\sqrt{a_0} \vee \varepsilon^{1/3}). \tag{3.12}$$

Proof: Let $c_0 \geqslant 1$ and $c_1 > T^2 y_{-T}/\varepsilon$ be constants to be chosen later, and denote by τ the first exit time of y_t from the strip $0 < y_t < c_1\varepsilon/t^2$. Set $t_0 = -c_0(\sqrt{a_0} \vee \varepsilon^{1/3})$. Then, for $-T \leqslant t \leqslant \tau \wedge t_0$, we get from (3.9) and (3.5) that

$$\frac{|b^\star(y,t)|}{yt^2} \leqslant M\frac{3x^\star(t) + y}{|t|}\frac{y}{|t|} \leqslant M'\biggl(1 + c_1\frac{\varepsilon}{|t|^3}\biggr)c_1\frac{\varepsilon}{|t|^3} \leqslant M'\biggl(1 + \frac{c_1}{c_0^3}\biggr)\frac{c_1}{c_0^3}, \tag{3.13}$$

for some constants $M, M' > 0$. The relations (3.8) and (3.10) yield the existence of constants $c_\pm > 0$ such that $a^\star(t) \leqslant -c_- t^2$ and $-\frac{\mathrm{d}}{\mathrm{d}t} x^\star(t) \leqslant c_+$ for $t \in [-T, t_0]$. From (3.7) and (3.13) we obtain

$$\varepsilon \frac{\mathrm{d}y}{\mathrm{d}t} \leqslant -c_- t^2 y \biggl[1 - \frac{M'}{c_-} \biggl(1 + \frac{c_1}{c_0^3} \biggr) \frac{c_1}{c_0^3} \biggr] + \varepsilon c_+. \tag{3.14}$$

For any given c_1, we can choose c_0 large enough for the term in brackets to be larger than 1/2. Then, by Lemma 3.2, there exists a constant $c_2 = c_2(c_+, c_-) > 0$ such that

$$y_t \leqslant y_{-T} \,\mathrm{e}^{-c_-(t^3+T^3)/6\varepsilon} + c_+ \int_{-T}^t \mathrm{e}^{-c_-(t^3-s^3)/6\varepsilon} \,\mathrm{d}s \leqslant c_2 \frac{\varepsilon}{t^2} \tag{3.15}$$

for all $t \in [-T, \tau \wedge t_0]$. Therefore, if $c_1 > c_2$, then $\tau \geqslant t_0$ follows.

The lower bound can be obtained in exactly the same way. □

For the remainder of this subsection, let $t_0 = -c_0(\sqrt{a_0} \vee \varepsilon^{1/3})$ with c_0 chosen according to the preceding proposition. Note that this proposition implies that $x_t \asymp x^\star(t) \asymp |t|$ for $-T \leqslant t \leqslant t_0$ and, in particular, that $y_{t_0} \asymp (\varepsilon/a_0) \wedge \varepsilon^{1/3}$.

We now consider the dynamics for $|t| \leqslant |t_0|$, starting with the case of a_0 not too small, i. e., the case of $y_{t_0} \asymp \varepsilon/a_0$.

Proposition 3.4. *There exists a constant $\gamma_0 > 0$, depending only on f and y_{t_0}, such that, when $a_0 \geqslant \gamma_0 \varepsilon^{2/3}$, then*

$$y_t = C_1(t)(t^\star - t) + C_2(t) \qquad \textit{with} \qquad C_1(t) \asymp \frac{\varepsilon}{a_0^{3/2}}, \quad C_2(t) \asymp \frac{\varepsilon^2}{a_0^{5/2}} \tag{3.16}$$

for all $|t| \leqslant |t_0|$.

Proof: Again, we will only show how to obtain an upper bound, since the corresponding lower bound can be established in exactly the same way.

First we fix a constant $c_1 > a_0 y_{t_0}/\varepsilon + 2(t^\star - t_0)/\sqrt{a_0} + 4\varepsilon/(c_- a_0^{3/2})$. We denote by τ the first exit time of y_t from the strip $|y_t| < c_1 \varepsilon / a_0$. For $t_0 \leqslant t \leqslant \tau \wedge |t_0|$, we have

$$\frac{|b^\star(y,t)|}{|y|} \leqslant M\bigl(3x^\star(t) + |y|\bigr)|y| \leqslant M' a_0 \biggl(1 + c_1 \frac{\varepsilon}{a_0^{3/2}} \biggr) c_1 \frac{\varepsilon}{a_0^{3/2}} \tag{3.17}$$

with constants $M, M' > 0$. Choosing γ_0 large enough, we get

$$\varepsilon \frac{\mathrm{d}y}{\mathrm{d}t} \leqslant -\frac{c_-}{2} a_0 y - c_- \frac{\varepsilon}{\sqrt{a_0}} (t - t^\star), \tag{3.18}$$

which implies

$$\begin{aligned} y_t &\leqslant y_{t_0} \,\mathrm{e}^{-c_- a_0 (t-t_0)/2\varepsilon} - \frac{c_-}{\sqrt{a_0}} \int_{t_0}^t \mathrm{e}^{-c_- a_0 (t-s)/2\varepsilon} (s - t^\star) \,\mathrm{d}s \\ &= \frac{2\varepsilon}{a_0^{3/2}} (t^\star - t) + \eta(\varepsilon) \,\mathrm{e}^{-c_- a_0 (t-t_0)/2\varepsilon} + \frac{4\varepsilon^2}{c_- a_0^{5/2}} \end{aligned} \tag{3.19}$$

by partial integration. Here

$$\eta(\varepsilon) = y_{t_0} - 2\frac{\varepsilon}{a_0}\biggl(\frac{t^\star - t_0}{\sqrt{a_0}} + \frac{2\varepsilon}{c_- a_0^{3/2}}\biggr) \tag{3.20}$$

satisfies $\eta(\varepsilon) = \mathcal{O}(\varepsilon/a_0)$.

We want to estimate the contribution of the middle term on the right-hand side of (3.19). Assume first that $\eta(\varepsilon) > 0$ and consider $t \leqslant t^\star$. By convexity,

$$\mathrm{e}^{-c_- a_0 (t-t_0)/2\varepsilon} \leqslant \frac{t^\star - t}{t^\star - t_0} + \mathrm{e}^{-c_- a_0 (t^\star - t_0)/2\varepsilon}. \tag{3.21}$$

Now,

$$\frac{2\varepsilon}{a_0^{3/2}} + \eta(\varepsilon)\frac{1}{t^\star - t_0} \asymp \frac{\varepsilon}{a_0^{3/2}}. \tag{3.22}$$

Since $X\,\mathrm{e}^{-X} \to 0$ as $X \to \infty$, we also have

$$\eta(\varepsilon)\,\mathrm{e}^{-c_- a_0 (t^\star - t_0)/2\varepsilon} + \frac{4\varepsilon^2}{c_- a_0^{5/2}} \asymp \frac{\varepsilon^2}{a_0^{5/2}}, \tag{3.23}$$

provided γ_0 is large enough. This shows the existence of constants $\bar C_1 \geqslant 2$ and $\bar C_2 > 0$ such that

$$y_t \leqslant \bar C_1 \frac{\varepsilon}{a_0^{3/2}}(t^\star - t) + \bar C_2 \frac{\varepsilon^2}{a_0^{5/2}} \qquad \text{for } t \leqslant t^\star \wedge \tau. \tag{3.24}$$

For $t \geqslant t^\star$,

$$\mathrm{e}^{-c_- a_0 (t-t_0)/2\varepsilon} \leqslant \mathrm{e}^{-c_- a_0 (t^\star - t_0)/2\varepsilon} \tag{3.25}$$

is immediate, and (3.23) shows that (3.24) also holds for $t^\star \leqslant t \leqslant \tau$. Note that in the case $\eta(\varepsilon) \leqslant 0$, (3.24) holds trivially. Since $y_t < c_1\varepsilon/a_0$ is a direct consequence of (3.19) and our choice of c_1, $\tau \geqslant |t_0|$ follows, and, therefore, the upper bound (3.24) holds for all $|t| \leqslant |t_0|$. □

Note that the result (3.16) implies that y_t changes sign at a time $t^\star + \mathcal{O}(\varepsilon/a_0)$, which shows that x_t actually crosses $x^\star(t)$ at a time $\tilde t$ satisfying $\tilde t - t^\star \asymp \varepsilon/a_0$. For large enough γ_0, the proposition also shows that $x_t \asymp \sqrt{a_0}$ for $|t| \leqslant |t_0|$ and that $y_{|t_0|} \asymp -\varepsilon/a_0$.

We consider now the case $a_0 < \gamma_0\varepsilon^{2/3}$ with γ_0 from Proposition 3.4. Without loss of generality, we may assume that $\gamma_0 \geqslant 1$.

Proposition 3.5. *Assume that $a_0 < \gamma_0\varepsilon^{2/3}$. Then, for any fixed $t_1 \asymp \varepsilon^{1/3}$,*

$$x_t \asymp \varepsilon^{1/3} \qquad \textit{for } t_0 \leqslant t \leqslant t_1, \tag{3.26}$$

and x_t crosses $x^\star(t)$ at a time $\tilde t$ satisfying $\tilde t \asymp \varepsilon^{1/3}$.

Proof: In order to show (3.26), we rescale space and time in the following way:

$$x = \varepsilon^{1/3} a_1^{1/6} z, \qquad t = \varepsilon^{1/3} a_1^{-1/3} s. \tag{3.27}$$

Let $s_0 = \varepsilon^{-1/3} a_1^{1/3} t_0$. Then $z_{s_0} \asymp 1$, and z satisfies the differential equation

$$\frac{\mathrm{d}z}{\mathrm{d}s} = \tilde a(s,\varepsilon) z + \bigl[-1 + r_0(\varepsilon^{1/3} a_1^{1/6} z, \varepsilon^{1/3} a_1^{-1/3} s)\bigr] z^3, \tag{3.28}$$

where

$$\tilde{a}(s,\varepsilon) = \frac{a(\varepsilon^{1/3}a_1^{-1/3}s)}{\varepsilon^{2/3}a_1^{1/3}} = \tilde{a}_0 + s^2 + \mathcal{O}(\varepsilon^{1/3}s^3), \qquad \tilde{a}_0 = \frac{a_0}{\varepsilon^{2/3}a_1^{1/3}} \leqslant \frac{\gamma_0}{a_1^{1/3}}. \tag{3.29}$$

(3.28) is a perturbation of order $\varepsilon^{1/3}$ of the Bernoulli equation

$$\frac{\mathrm{d}z}{\mathrm{d}s} = \tilde{a}(s)z - z^3, \qquad \tilde{a}(s) = \tilde{a}_0 + s^2. \tag{3.30}$$

Using Gronwall's inequality, one easily shows that on an s-time scale of order 1, the solution of (3.28) differs by $\mathcal{O}(\varepsilon^{1/3})$ from the solution of (3.30), which is

$$z_s = \frac{z_{s_0}\,\mathrm{e}^{\tilde{\alpha}(s,s_0)}}{\left(1 + 2z_{s_0}^2\displaystyle\int_{s_0}^{s}\mathrm{e}^{2\tilde{\alpha}(u,s_0)}\,\mathrm{d}u\right)^{1/2}}, \qquad \text{where } \tilde{\alpha}(s,s_0) = \int_{s_0}^{s}\tilde{a}(u)\,\mathrm{d}u. \tag{3.31}$$

This function is bounded away from zero, and remains of order one for s of order one, which shows that $x_t \asymp \varepsilon^{1/3}$ on $[t_0, t_1]$.

Since $x_t \asymp \varepsilon^{1/3}$ and $x^\star(t) \asymp \sqrt{a_0} \vee |t|$, x_t and $x^\star(t)$ necessarily cross at some time $\tilde{t} \asymp \varepsilon^{1/3}$. □

Note that the above proposition also implies bounds on y_t, namely, $y_t = \mathcal{O}(\varepsilon^{1/3})$ for $t_0 \leqslant t \leqslant t_1$, and there exist constants $\tilde{c}_+ > \tilde{c}_- > 0$ such that

$$y_t \asymp \begin{cases} \varepsilon^{1/3} & \text{for } t_0 \leqslant t \leqslant \tilde{c}_-\varepsilon^{1/3}, \\ 0 & \text{for } t = \tilde{t}, \\ -\varepsilon^{1/3} & \text{for } \tilde{c}_+\varepsilon^{1/3} \leqslant t \leqslant t_1. \end{cases} \tag{3.32}$$

Gathering the results for $a_0 \geqslant \gamma_0\varepsilon^{2/3}$ and $a_0 < \gamma_0\varepsilon^{2/3}$, we see that there exists a time $t_1 \asymp (\sqrt{a_0} \vee \varepsilon^{1/3})$ such that $y_{t_1} \asymp -\varepsilon/t_1^2$. By enlarging c_0 if necessary, we may assume that $t_1 = c_0(\sqrt{a_0} \vee \varepsilon^{1/3})$.

Proposition 3.6. *On the interval* $[t_1, T]$,

$$y_t \asymp -\frac{\varepsilon}{t^2}. \tag{3.33}$$

Proof: The proof is similar to the one of Proposition 3.3. □

Note that the previous result implies $x_t \asymp x^\star(t) \asymp t$, provided c_0 is large enough.

So far, we have proved that for $t \in [-T, T]$, x_t tracks $x^\star(t)$ at a distance of order

$$\frac{\varepsilon}{t^2} \wedge \frac{\varepsilon}{a_0} \wedge \varepsilon^{1/3}, \tag{3.34}$$

and that the two curves cross at a time $\tilde{t}$ satisfying $\tilde{t} - t^\star \asymp (\varepsilon/a_0) \wedge \varepsilon^{1/3}$. Let us now examine the behaviour of the linearization

$$\bar{a}(t) = \partial_x f(x_t, t), \tag{3.35}$$

which will determine the behaviour of orbits starting close to the particular solution x_t.

Proposition 3.7. *For all $t \in [-T,T]$ and all $a_0 = \mathcal{O}_\varepsilon(1)$,*

$$\bar{a}(t) \asymp -(t^2 \vee a_0 \vee \varepsilon^{2/3}). \tag{3.36}$$

PROOF: By Taylor's formula we get

$$\begin{aligned}\bar{a}(t) &= \partial_x f(x^\star(t) + y_t, t) \\ &= a^\star(t) - \Big[6 + \mathcal{O}\big(x^\star(t) + y_t\big)\Big]\Big(x^\star(t) + \frac{1}{2}y_t\Big)y_t.\end{aligned} \tag{3.37}$$

Consider first the case $a_0 \geqslant \gamma_0\varepsilon^{2/3}$ for γ_0 large enough. Equation (3.8) implies that $a^\star(t) \leqslant -c_-(a_0 \vee t^2)$ for a constant $c_- > 0$. On the other hand, (3.34) shows that the second term on the right-hand side of (3.37) is bounded in absolute value by $c_+\varepsilon/a_0$ for a constant $c_+ > 0$. Thus if $\gamma_0 > (c_+/c_-)^{2/3}$ we obtain that $\bar{a}(t) \asymp a^\star(t) \asymp -(a_0 \vee t^2)$.

We consider next the case $a_0 < \gamma_0\varepsilon^{2/3}$. For $|t| \geqslant c_0\varepsilon^{1/3}$, the above argument can be repeated. The non-trivial case occurs for $|t| < c_0\varepsilon^{1/3}$. By rescaling variables as in Proposition 3.5, we obtain that

$$\bar{a}(t) = \varepsilon^{2/3}a_1^{1/3}\big[\tilde{a}(s,\varepsilon) - 3z_s^2 + \mathcal{O}(\varepsilon^{1/3})\big]. \tag{3.38}$$

We have to show that $\bar{a}(t) \asymp -\varepsilon^{2/3}$ which is equivalent to $\tilde{a}(s,\varepsilon) - 3z_s^2 \asymp -1$ for s of order one. The lower bound is trivial as $\tilde{a}(s,\varepsilon) \geqslant 0$ and $z_s \asymp 1$. In order to show the upper bound, first note that for $t \leqslant 0$, we have $x_t > x^\star(t)$ which implies $\bar{a}(t) < a^\star(t)$. Therefore, it is sufficient to consider $s \geqslant 0$. Taking into account the expression (3.31), we find that showing the upper bound amounts to showing that

$$\Big(const + \frac{\tilde{a}(s)}{z_{s_0}^2}\Big)\,\mathrm{e}^{-2\tilde{\alpha}(s,s_0)} + 2\tilde{a}(s)\int_{s_0}^{s} \mathrm{e}^{-2\tilde{\alpha}(s,u)}\,\mathrm{d}u < 3. \tag{3.39}$$

Since $|s_0|$ is proportional to c_0, choosing *a priori* a large enough c_0 also makes $|s_0|$ large. Thus it is in fact sufficient to verify that

$$2\tilde{a}(s)\int_{-\infty}^{s} \mathrm{e}^{-2\tilde{\alpha}(s,u)}\,\mathrm{d}u < 3 \tag{3.40}$$

for all $s \geqslant 0$. Optimizing the left-hand side with respect to $\tilde{a}_0 \geqslant 0$ and s shows that we may assume $\tilde{a}_0 = 0$ and that (3.40) holds. □

3.2 The random motion near the stable equilibrium branches

We now consider the SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon}f(x_t,t)\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad x_{-T} = x_0, \tag{3.41}$$

on the time interval $[-T,T]$, where we assume $x_0 - x^\star(-T) \asymp \varepsilon$. In order to compare the solution x_t with the solution x_t^{det} of the corresponding deterministic equation (3.6), we introduce the difference $y_t = x_t - x_t^{\mathrm{det}}$, which satisfies the SDE

$$\mathrm{d}y_t = \frac{1}{\varepsilon}\big[\bar{a}(t)y_t + \bar{b}(y_t,t)\big]\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad y_{-T} = 0, \tag{3.42}$$

where $\bar{a}(t)$ is the linearization (3.35) of f along $x_t^{\det}$, and Taylor's formula yields the relations

$$\begin{aligned} \bar{b}(y,t) &= -\big[1+\mathcal{O}(x_t^{\det}+|y|)+\mathcal{O}(t)\big](3x_t^{\det}+y)y^2, \\ |\bar{b}(y,t)| &\leqslant M(x_t^{\det}+|y|)y^2 \end{aligned} \tag{3.43}$$

whenever $|t| \leqslant T$ and $x_t^{\det}+|y| \leqslant d$, where M is a positive constant. Let us first consider the linearization of (3.42), namely

$$\mathrm{d}y_t^0 = \frac{1}{\varepsilon}\bar{a}(t)y_t^0\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad y_{-T}^0 = 0. \tag{3.44}$$

The random variable y_t^0 is Gaussian with expectation zero and variance

$$v(t) = \frac{\sigma^2}{\varepsilon}\int_{-T}^t \mathrm{e}^{2\overline{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s, \qquad \text{where } \overline{\alpha}(t,s) = \int_s^t \bar{a}(u)\,\mathrm{d}u. \tag{3.45}$$

Lemma 3.2 and Proposition 3.7 imply that

$$\zeta(t) := \frac{1}{2|\bar{a}(-T)|}\,\mathrm{e}^{2\overline{\alpha}(t,-T)/\varepsilon} + \frac{1}{\varepsilon}\int_{-T}^t \mathrm{e}^{2\overline{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s \asymp \frac{1}{t^2 \vee a_0 \vee \varepsilon^{2/3}}. \tag{3.46}$$

Thus $v(t)$ is of order $\sigma^2/(t^2 \vee a_0 \vee \varepsilon^{2/3})$, except for t very close to $-T$. We now show that y_t^0 is likely to remain in a strip of width proportional to $\sqrt{\zeta(t)}$.

Proposition 3.8. *For $-T \leqslant t \leqslant T$,*

$$\mathbb{P}^{-T,0}\Big\{\sup_{-T\leqslant s\leqslant t}\frac{|y_s^0|}{\sqrt{\zeta(s)}} \geqslant h\Big\} \leqslant C(t,\varepsilon)\exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\big(1-\mathcal{O}(\varepsilon)\big)\Big\}, \tag{3.47}$$

where

$$C(t,\varepsilon) = \frac{|\overline{\alpha}(t,-T)|}{\varepsilon^2} + 2. \tag{3.48}$$

Proof: Let $-T = u_0 < u_1 < \dots < u_K = t$, with some $K > 0$, be a partition of $[-T,t]$. In [BG, Lemma 3.2], we show that the probability (3.47) is bounded above by

$$2\sum_{k=1}^K P_k, \qquad \text{where } P_k = \exp\Big\{-\frac{1}{2}\frac{h^2}{\sigma^2}\inf_{u_{k-1}\leqslant s\leqslant u_k}\frac{\zeta(s)}{\zeta(u_k)}\,\mathrm{e}^{2\overline{\alpha}(u_k,s)/\varepsilon}\Big\}. \tag{3.49}$$

Now we choose the partition by requiring that

$$\overline{\alpha}(u_k,u_{k-1}) = -2\varepsilon^2 \qquad \text{for } 1 \leqslant k < K = \bigg\lceil\frac{|\overline{\alpha}(t,-T)|}{2\varepsilon^2}\bigg\rceil. \tag{3.50}$$

Since $\bar{a}(s) < 0$, we have $\zeta'(s) = [2\bar{a}(s)\zeta(s)+1]/\varepsilon \leqslant 1/\varepsilon$, and thus

$$\inf_{u_{k-1}\leqslant s\leqslant u_k}\frac{\zeta(s)}{\zeta(u_k)} \geqslant \frac{1}{\zeta(u_k)}\inf_{u_{k-1}\leqslant s\leqslant u_k}\Big[\zeta(u_k) - \frac{u_k - s}{\varepsilon}\Big] = 1 - \frac{u_k - u_{k-1}}{\varepsilon\zeta(u_k)}. \tag{3.51}$$

If k is such that $|u_k| \geqslant \sqrt{a_0} \vee \varepsilon^{1/3}$, then by (3.50) and (3.36), there is a constant c_- such that

$$2\varepsilon^2 \geqslant c_- \int_{u_{k-1}}^{u_k} s^2 \, \mathrm{d}s \geqslant \frac{c_-}{6} u_k^2 (u_k - u_{k-1}), \tag{3.52}$$

and hence by (3.46) (choosing the same c_- for brevity of notation)

$$\frac{u_k - u_{k-1}}{\zeta(u_k)} \leqslant \frac{12\varepsilon^2}{c_- u_k^2} \frac{u_k^2}{c_-} = \mathcal{O}(\varepsilon^2). \tag{3.53}$$

For all other k, we have

$$2\varepsilon^2 \geqslant c_- (a_0 \vee \varepsilon^{2/3})(u_k - u_{k-1}) \qquad \Rightarrow \qquad \frac{u_k - u_{k-1}}{\zeta(u_k)} \leqslant \frac{2\varepsilon^2}{c_-^2}. \tag{3.54}$$

In both cases, we find

$$P_k \leqslant \exp\Bigl\{-\frac{1}{2}\frac{h^2}{\sigma^2}\bigl(1 - \mathcal{O}(\varepsilon)\bigr)\Bigr\}, \tag{3.55}$$

which leads to the result, using the definition of K. $\square$

Let us now compare solutions of the linear equation (3.44) and the nonlinear equation (3.42). We introduce the events

$$\Omega_t(h) = \bigl\{\omega\colon |y_s| \leqslant h\sqrt{\zeta(s)} \;\; \forall s \in [-T, t]\bigr\} \tag{3.56}$$

$$\Omega_t^0(h) = \bigl\{\omega\colon |y_s^0| \leqslant h\sqrt{\zeta(s)} \;\; \forall s \in [-T, t]\bigr\}. \tag{3.57}$$

Notation 3.9. *For two events Ω_1 and Ω_2, we write $\Omega_1 \overset{\text{a.s.}}{\subset} \Omega_2$ if $\mathbb{P}$-almost all $\omega \in \Omega_1$ belong to Ω_2.*

Proposition 3.10. *There exists a constant ϱ, depending only on f, such that*

- *if $-T \leqslant t \leqslant -(\sqrt{a_0} \vee \varepsilon^{1/3})$ and $h < t^2/\varrho$, then*

$$\Omega_t^0(h) \overset{\text{a.s.}}{\subset} \Omega_t\Bigl(\Bigl[1 + \varrho\frac{h}{t^2}\Bigr]h\Bigr); \tag{3.58}$$

- *if $-(\sqrt{a_0} \vee \varepsilon^{1/3}) \leqslant t \leqslant T$ and $h < (a_0 \vee \varepsilon^{2/3})/\varrho$, then*

$$\Omega_t^0(h) \overset{\text{a.s.}}{\subset} \Omega_t\Bigl(\Bigl[1 + \varrho\frac{h}{a_0 \vee \varepsilon^{2/3}}\Bigr]h\Bigr). \tag{3.59}$$

Proof: The proof is based on the fact that the variable $z_s = y_s - y_s^0$ satisfies the relation

$$z_s = \frac{1}{\varepsilon}\int_{-T}^{s} \mathrm{e}^{\overline{\alpha}(s,u)/\varepsilon}\, \bar{b}(y_u, u)\, \mathrm{d}u. \tag{3.60}$$

Consider first the case $-T \leqslant t \leqslant -(\sqrt{a_0} \vee \varepsilon^{1/3})$. Let $\varrho > 0$ be a constant to be chosen later, and set $\delta = \varrho h/t^2 < 1$. We define the first exit time

$$\tau = \inf\bigl\{s \in [-T, t]\colon |z_s| \geqslant \delta h\sqrt{\zeta(s)}\bigr\} \in [-T, t] \cup \{\infty\}. \tag{3.61}$$

Pick any $\omega \in A := \Omega_t^0(h) \cap \{\omega\colon \tau(\omega) < \infty\}$ and $s \in [-T, \tau(\omega)]$. Then we have

$$|y_u^0(\omega)| \leqslant h\sqrt{\zeta(u)}, \qquad |y_u(\omega)| \leqslant (1+\delta)h\sqrt{\zeta(u)} < 2h\sqrt{\zeta(u)} \tag{3.62}$$

for all $u \in [-T, s]$. From (3.5) and (3.46), we obtain the existence of a constant $c_+ > 0$ such that

$$x_u^{\det} \leqslant c_+|u|, \qquad |y_u(\omega)| < 2h\frac{\sqrt{c_+}}{|u|} \tag{3.63}$$

for these u. Hence, by (3.43) we get the estimate

$$|\bar{b}(y_u, u)| < M\Big(c_+|u| + 2h\frac{\sqrt{c_+}}{|u|}\Big)4h^2\frac{c_+}{u^2} \leqslant 4M\frac{h^2 c_+^2}{|s|}\Big(1 + \frac{2h}{\sqrt{c_+}s^2}\Big) \tag{3.64}$$

and thus, by (3.60) and Lemma 3.2,

$$|z_s| < 4M\frac{h^2 c_+^2}{|s|}\Big(1 + \frac{2h}{\sqrt{c_+}s^2}\Big)\frac{1}{\varepsilon}\int_{-T}^{s} \mathrm{e}^{\overline{\alpha}(s,u)/\varepsilon}\,\mathrm{d}u \leqslant 4M\frac{h^2 c_+^3}{|s|^3}\Big(1 + \frac{2h}{\sqrt{c_+}s^2}\Big), \tag{3.65}$$

where we use again the same c_+ for brevity of notation. Using (3.46) once again, we arrive at the bound

$$\frac{|z_s|}{h\sqrt{\zeta(s)}} < 4M\frac{c_+^3}{\sqrt{c_-}}\frac{h}{s^2}\Big(1 + \frac{2}{\sqrt{c_+}}\frac{h}{s^2}\Big). \tag{3.66}$$

Now we choose

$$\varrho = \frac{2}{\sqrt{c_+}} \vee 8M\frac{c_+^3}{\sqrt{c_-}}, \tag{3.67}$$

which implies

$$\frac{|z_s|}{h\sqrt{\zeta(s)}} < \frac{\varrho}{2}\frac{h}{s^2}\Big(1 + \varrho\frac{h}{s^2}\Big) \leqslant \frac{\delta}{2}(1+\delta) < \delta \tag{3.68}$$

for all $s \in [-T, \tau(\omega)]$, by the definition of δ. Hence $|z_{\tau(\omega)}| < \delta h\sqrt{\zeta(\tau(\omega))}$ for almost all $\omega \in A$. Since we have $|z_{\tau(\omega)}| = \delta h\sqrt{\zeta(\tau(\omega))}$ whenever $\tau(\omega) < \infty$, we conclude that $\mathbb{P}(A) = 0$, and thus $\tau(\omega) = \infty$ for almost all $\omega \in \Omega_t^0(h)$, which implies that $|y_s(\omega)| \leqslant (1+\delta)h\sqrt{\zeta(s)}$ for $-T \leqslant s \leqslant t$ and these ω. This completes the proof of (3.58).

The proof of (3.59) is almost the same. In the case $-(\sqrt{a_0} \vee \varepsilon^{1/3}) \leqslant t \leqslant T$, we take $\delta = \varrho h/(a_0 \vee \varepsilon^{2/3})$. The estimate (3.63) has to be replaced by

$$x_u^{\det} \leqslant c_+(|u| \vee \sqrt{a_0} \vee \varepsilon^{1/3}), \qquad |y_u(\omega)| \leqslant 2h\frac{\sqrt{c_+}}{|u| \vee \sqrt{a_0} \vee \varepsilon^{1/3}}, \tag{3.69}$$

and thus we get, instead of (3.65), the bound

$$|z_s| \leqslant 4M\frac{h^2 c_+^3}{(\sqrt{a_0} \vee \varepsilon^{1/3})(s^2 \vee a_0 \vee \varepsilon^{2/3})}\Big(1 + \frac{2h}{\sqrt{c_+}(a_0 \vee \varepsilon^{2/3})}\Big). \tag{3.70}$$

The remainder of the proof is similar. □

Now, the preceding two propositions immediately imply Theorem 2.2, as Proposition 3.8 shows the desired behaviour for the approximation by a Gaussian process and Proposition 3.10 allows to extend this result to the original process.

3.3 The transition regime

We consider now the regime of σ sufficiently large to allow for transitions from one stable equilibrium branch to the other. Here $x_t^{\det}$ is the solution of the deterministic equation (3.6) with the same initial condition $x_{-T}^{\det}$ as in the previous sections, which tracks $x^\star(t)$ at distance at most $\mathcal{O}(\varepsilon^{1/3})$. x_t denotes a general solution of the SDE (3.1). Our aim is to establish an upper bound for the probability of *not* reaching the axis $x = 0$, which, by using symmetry, will allow us to estimate the transition probability. Let $\delta > 0$ be the constant defined in (2.23), i.e. by

$$x\partial_{xx}f(x,t) \leqslant 0 \qquad \text{for } |x| \leqslant \delta \text{ and } |t| \leqslant T. \tag{3.71}$$

The basic ingredient of our estimate is the following comparison lemma which allows us to linearize the stochastic differential equations under consideration and, therefore, to investigate Gaussian approximations to our processes. The lemma gives conditions under which relations between initial conditions carry over to the sample paths.

Lemma 3.11. *Fix some initial time $t_0 \in [-T,T]$. We consider the following processes on $[t_0,T]$:*

- *the solution $x_t^{\det}$ of the deterministic differential equation* (3.6) *with initial condition $x_{t_0}^{\det} \in [0,\delta]$;*
- *the solution x_t of the SDE* (3.41) *with initial condition $x_{t_0} \in [x_{t_0}^{\det},\delta]$;*
- *the difference $y_t = x_t - x_t^{\det}$, which satisfies $y_{t_0} = x_{t_0} - x_{t_0}^{\det} \geqslant 0$;*
- *the solution y_t^0 of the linear SDE*

$$\mathrm{d}y_t^0 = \frac{1}{\varepsilon}\tilde{a}(t)y_t^0\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad \textit{where } \tilde{a}(t) = \partial_x f(x_t^{\det},t) \tag{3.72}$$

with initial condition $y_{t_0}^0 \in [y_{t_0}, \delta - x_{t_0}^{\det}]$.

If $0 \leqslant y_s^0 \leqslant \delta - x_s^{\det}$ for all $s \in [t_0,t]$, then $y_s \leqslant y_s^0$ for those s. Similarly, if $0 \leqslant y_s \leqslant \delta - x_s^{\det}$ for all $s \in [t_0,t]$, then $y_s^0 \geqslant y_s$ for those s. The result remains true when t is replaced by a stopping time.

Proof: The hypothesis (3.71) implies that for all $y \in [0, \delta - x_s^{\det}]$,

$$f(x_s^{\det} + y, s) \leqslant f(x_s^{\det}, s) + \tilde{a}(s)y. \tag{3.73}$$

Let $\tau = \inf\{s \in [t_0,t] \colon y_s \notin [0,\delta - x_s^{\det}]\} \in [t_0,t] \cup \{\infty\}$. For $t_0 \leqslant s \leqslant \tau$, the variable $z_s = y_s - y_s^0$ satisfies

$$\begin{aligned} z_s &= z_{t_0} + \frac{1}{\varepsilon}\int_{t_0}^{s}\bigl[f(x_u^{\det}+y_u,u) - f(x_u^{\det},u) - \tilde{a}(u)y_u^0\bigr]\,\mathrm{d}u \\ &\leqslant z_{t_0} + \frac{1}{\varepsilon}\int_{t_0}^{s}\tilde{a}(u)z_u\,\mathrm{d}u, \qquad z_{t_0} \leqslant 0. \end{aligned} \tag{3.74}$$

Applying Gronwall's inequality, we obtain

$$z_s \leqslant z_{t_0}\,\mathrm{e}^{\tilde{\alpha}(s,t_0)/\varepsilon} \leqslant 0 \qquad \forall s \in [t_0, \tau \wedge t], \tag{3.75}$$

where $\tilde{\alpha}(s,t_0) = \int_{t_0}^{s}\tilde{a}(u)\,\mathrm{d}u$. This proves the result for $t_0 \leqslant s \leqslant \tau\wedge t$. Now if y_s is negative, the result is trivially satisfied, and if y_s becomes positive again, the above argument can be repeated. Note that $y_\tau \leqslant \delta - x_\tau^{\det}$ is immediate. This proves the first assertion, and the second assertion can be proved directly, without use of τ. $\square$

We will now proceed as follows. Let $\zeta(t)$ be the function defined in (3.46), and let h be such that $x_s^{\det} + h\sqrt{\zeta(s)} < \delta$ for all $s \in [t_0, t]$. Given $x_0 \in (0, \delta)$, we can write

$$\begin{aligned}
\mathbb{P}^{t_0,x_0}\bigl\{x_s > 0 \ \ \forall s \in [t_0,t]\bigr\} \leqslant \mathbb{P}^{t_0,x_0}\Bigl\{\sup_{t_0\leqslant s\leqslant t} \frac{x_s - x_s^{\det}}{\sqrt{\zeta(s)}} > h\Bigr\} \\
+ \mathbb{P}^{t_0,x_0}\bigl\{0 < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \ \ \forall s \in [t_0,t]\bigr\}. \qquad (3.76)
\end{aligned}$$

We will estimate these two terms separately. The first event is similar to the event we have examined in the previous subsection, but we need here an estimate valid for all times, even when σ is not very small, whereas the previous result is only useful for $\sigma \leqslant t^2 \vee a_0 \vee \varepsilon^{2/3}$. We will show the following.

Proposition 3.12. *Assume* $0 \leqslant x_0 \leqslant x_{t_0}^{\det} + \frac12 h\sqrt{\zeta(t_0)}$. *Then*

$$\mathbb{P}^{t_0,x_0}\Bigl\{\sup_{t_0\leqslant s\leqslant t} \frac{x_s - x_s^{\det}}{\sqrt{\zeta(s)}} > h\Bigr\} \leqslant \frac52 \Bigl(\frac{|\overline{\alpha}(t,t_0)|}{\varepsilon} + 1\Bigr) \mathrm{e}^{-\bar\kappa h^2/\sigma^2}, \qquad (3.77)$$

where $\bar\kappa$ *is a positive constant and* $\overline{\alpha}(t,t_0) = \int_{t_0}^t \bar a(s)\,\mathrm{d}s$.

Proof:

1. We define a partition $t_0 = u_0 < u_1 < \dots < u_K = t$ of the interval $[t_0,t]$ by requiring

$$|\overline{\alpha}(u_k,u_{k-1})| = \varepsilon \qquad \text{for } 1 \leqslant k < K = \biggl\lceil \frac{|\overline{\alpha}(t,t_0)|}{\varepsilon} \biggr\rceil. \qquad (3.78)$$

Note that similar arguments as in the proof of Proposition 3.8 yield

$$\frac{u_{k+1} - u_k}{\zeta(u_k)} = \mathcal{O}(\varepsilon) \qquad \text{for all } k. \qquad (3.79)$$

Now let $\rho_k = \frac12 h\sqrt{\zeta(u_k)}$ and $y_s = x_s - x_s^{\det}$ as usual. Define

$$\begin{aligned}
Q_k = \sup_{y_k\leqslant\rho_k} \biggl[\mathbb{P}^{u_k,y_k}\Bigl\{\sup_{u_k\leqslant s\leqslant u_{k+1}} \frac{y_s}{\sqrt{\zeta(s)}} > h\Bigr\} \\
+ \mathbb{P}^{u_k,y_k}\Bigl\{\sup_{u_k\leqslant s\leqslant u_{k+1}} \frac{y_s}{\sqrt{\zeta(s)}} \leqslant h,\ y_{u_{k+1}} > \rho_{k+1}\Bigr\}\biggr], \qquad (3.80)
\end{aligned}$$

for $0 \leqslant k < K-1$, and

$$Q_{K-1} = \sup_{y_{K-1}\leqslant\rho_{K-1}} \mathbb{P}^{u_{K-1},y_{K-1}}\Bigl\{\sup_{u_{K-1}\leqslant s\leqslant u_K} \frac{y_s}{\sqrt{\zeta(s)}} > h\Bigr\}. \qquad (3.81)$$

Then

$$\begin{aligned}
&\mathbb{P}^{t_0,x_0}\Bigl\{\sup_{t_0\leqslant s\leqslant t} \frac{x_s - x_s^{\det}}{\sqrt{\zeta(s)}} > h\Bigr\} \\
&\qquad \leqslant \mathbb{P}^{t_0,y_{t_0}}\Bigl\{\sup_{t_0\leqslant s\leqslant u_1} \frac{y_s}{\sqrt{\zeta(s)}} > h\Bigr\} + \mathbb{P}^{t_0,y_{t_0}}\Bigl\{\sup_{t_0\leqslant s\leqslant u_1} \frac{y_s}{\sqrt{\zeta(s)}} \leqslant h,\ y_{u_1} > \rho_1\Bigr\} \\
&\qquad\quad + \mathbb{E}^{t_0,y_{t_0}}\Bigl\{1_{\{y_{u_1}\leqslant\rho_1\}} \mathbb{P}^{u_1,y_{u_1}}\Bigl\{\sup_{u_1\leqslant s\leqslant t} \frac{y_s}{\sqrt{\zeta(s)}} > h\Bigr\}\Bigr\} \\
&\qquad \leqslant \dots \leqslant \sum_{k=0}^{K-1} Q_k. \qquad (3.82)
\end{aligned}$$

2. In order to estimate Q_k, we introduce the stochastic process $(y_s^{(k)})_{s\in[u_k,u_{k+1}]}$ defined by

$$y_s^{(k)} = \rho_k \, \mathrm{e}^{\overline{\alpha}(s,u_k)/\varepsilon} + \frac{\sigma}{\sqrt{\varepsilon}} \int_{u_k}^{s} \mathrm{e}^{\overline{\alpha}(s,u)/\varepsilon} \, \mathrm{d}W_u^{(k)}, \tag{3.83}$$

where $(W_u^{(k)})_{u\in[u_k,u_{k+1}]}$ is the Brownian motion $W_u^{(k)} = W_u - W_{u_k}$. Note that $y^{(k)}$ is the solution of the SDE (3.72) with initial condition $y_{u_k}^{(k)} = \rho_k$ at time u_k. We define the stopping times

$$\begin{aligned} \tau^0 &= \inf\bigl\{s \in [u_k, u_{k+1}] \colon y_s^{(k)} = 0\bigr\} \\ \tau^+ &= \inf\bigl\{s \in [u_k, u_{k+1}] \colon y_s^{(k)} = h\sqrt{\zeta(s)}\,\bigr\} \end{aligned} \tag{3.84}$$

describing the time when $y_s^{(k)}$ either reaches the t-axis or the upper boundary $h\sqrt{\zeta(s)}$. Now, Lemma 3.11 implies that if $y_{u_k} \leqslant \rho_k$, then $y_s \leqslant y_s^{(k)}$ for $u_k \leqslant s \leqslant \tau^0 \wedge \tau^+$. This shows that

$$Q_k \leqslant \mathbb{P}^{u_k,\rho_k}\bigl\{\tau^0 < u_{k+1}\bigr\} + \mathbb{P}^{u_k,\rho_k}\bigl\{\tau^+ < u_{k+1}\bigr\} + \mathbb{P}^{u_k,\rho_k}\bigl\{y_{u_{k+1}}^{(k)} > \rho_{k+1}\bigr\} \tag{3.85}$$

for $0 \leqslant k < K-1$, and

$$Q_{K-1} \leqslant \mathbb{P}^{u_{K-1},\rho_{K-1}}\bigl\{\tau^0 < u_K\bigr\} + \mathbb{P}^{u_{K-1},\rho_{K-1}}\bigl\{\tau^+ < u_K\bigr\}. \tag{3.86}$$

Each of these terms depends only on $y^{(k)}$, and can be easily estimated. Let

$$v_{u_{k+1}}^{(k)} = \frac{\sigma^2}{\varepsilon} \int_{u_k}^{u_{k+1}} \mathrm{e}^{2\overline{\alpha}(u_{k+1},u)/\varepsilon} \, \mathrm{d}u \tag{3.87}$$

denote the variance of $y_{u_{k+1}}^{(k)}$. Then by symmetry (as in (2.22)), we have

$$\begin{aligned} \mathbb{P}^{u_k,\rho_k}\bigl\{\tau^0 < u_{k+1}\bigr\} &= 2\mathbb{P}^{u_k,\rho_k}\bigl\{y_{u_{k+1}}^{(k)} < 0\bigr\} \\ &= \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{-\rho_k \, \mathrm{e}^{\overline{\alpha}(u_{k+1},u_k)/\varepsilon} \left(v_{u_{k+1}}^{(k)}\right)^{-1/2}} \mathrm{e}^{-z^2/2} \, \mathrm{d}z \\ &\leqslant \exp\biggl\{-\frac{1}{2} \frac{\rho_k^2 \, \mathrm{e}^{2\overline{\alpha}(u_{k+1},u_k)/\varepsilon}}{v_{u_{k+1}}^{(k)}}\biggr\} \\ &\leqslant \exp\biggl\{-\frac{1}{8} \, \mathrm{e}^{-2} \frac{h^2}{\sigma^2} \frac{\zeta(u_k)}{v_{u_{k+1}}^{(k)}/\sigma^2}\biggr\}. \end{aligned} \tag{3.88}$$

The second term on the right-hand side of (3.85) or (3.86), respectively, can be estimated using the symmetry (in distribution) of (3.83) under the map $\sigma \mapsto -\sigma$:

$$\begin{aligned} \mathbb{P}^{u_k,\rho_k}\bigl\{\tau^+ < u_{k+1}\bigr\} &= \mathbb{P}^{u_k,\rho_k}\bigl\{\exists s \in [u_k, u_{k+1}] \colon y_s^{(k)} \geqslant h\sqrt{\zeta(s)}\bigr\} \\ &\leqslant \mathbb{P}^{u_k,\rho_k}\bigl\{\exists s \in [u_k, u_{k+1}] \colon y_s^{(k)} \leqslant h\sqrt{\zeta(u_k)} \, \mathrm{e}^{\overline{\alpha}(s,u_k)/\varepsilon} - h\sqrt{\zeta(s)}\bigr\} \\ &\leqslant \mathbb{P}^{u_k,\rho_k}\bigl\{\tau^0 < u_{k+1}\bigr\}. \end{aligned} \tag{3.89}$$

In order to estimate the third term on the right-hand side of (3.85), we will use the fact that for $k < K-1$

$$\begin{aligned}
\zeta(u_{k+1}) &= \zeta(u_k)\,\mathrm{e}^{2\overline{\alpha}(u_{k+1},u_k)/\varepsilon} + \frac{1}{\varepsilon}\int_{u_k}^{u_{k+1}} \mathrm{e}^{2\overline{\alpha}(u_{k+1},s)/\varepsilon}\,\mathrm{d}s \\
&\geqslant \zeta(u_k)\,\mathrm{e}^{-2} + \frac{1-\mathrm{e}^{-2}}{2} \inf_{u_k\leqslant s\leqslant u_{k+1}} \frac{1}{|\bar{a}(s)|}.
\end{aligned} \tag{3.90}$$

Proposition 3.7 and (3.46) thus yield the existence of a constant $c_- > 0$ such that

$$\frac{\zeta(u_{k+1})}{\zeta(u_k)} \geqslant \mathrm{e}^{-2} + \frac{1-\mathrm{e}^{-2}}{2c_-^2}. \tag{3.91}$$

This allows us to estimate (for $k < K-1$)

$$\begin{aligned}
\mathbb{P}^{u_k,\rho_k}\bigl\{y^{(k)}_{u_{k+1}} > \rho_{k+1}\bigr\} &\leqslant \frac{1}{2}\exp\biggl\{-\frac{1}{2}\frac{(\rho_{k+1}-\rho_k\,\mathrm{e}^{\overline{\alpha}(u_{k+1},u_k)/\varepsilon})^2}{v^{(k)}_{u_{k+1}}}\biggr\} \\
&= \frac{1}{2}\exp\biggl\{-\frac{h^2}{8}\frac{\zeta(u_k)}{v^{(k)}_{u_{k+1}}}\biggl(\sqrt{\frac{\zeta(u_{k+1})}{\zeta(u_k)}} - \mathrm{e}^{-1}\biggr)^2\biggr\}.
\end{aligned} \tag{3.92}$$

3. The estimates (3.88), (3.89) and (3.92), inserted in (3.85) and (3.86), imply that

$$Q_k \leqslant \frac{5}{2}\exp\Bigl\{-\kappa_k\frac{h^2}{\sigma^2}\Bigr\}, \tag{3.93}$$

with

$$\kappa_k = \frac{1}{8}\frac{\zeta(u_k)}{v^{(k)}_{u_{k+1}}/\sigma^2}\,\mathrm{e}^{-2}\Biggl[1 \wedge \biggl(\sqrt{1+\frac{\mathrm{e}^2-1}{2c_-^2}} - 1\biggr)^2\Biggr]. \tag{3.94}$$

By (3.87), for each k, there exists a $\theta_k \in [\mathrm{e}^{-2}, 1]$ such that

$$\frac{v^{(k)}_{u_{k+1}}}{\sigma^2} = \frac{(u_{k+1}-u_k)}{\varepsilon}\theta_k. \tag{3.95}$$

Together with (3.79), this implies that $\kappa_k \asymp 1$ for all k, and thus the result follows from (3.82) with $\bar{\kappa} = \inf_k \kappa_k$. □

We now give an estimate of the second term in (3.76). The Markov property implies that we will obtain an upper bound by starting at time $-c_1\sqrt{\sigma}$.

Proposition 3.13. *There exist constants $c_1 > 0$ and $\bar{\kappa} > 0$ such that, if $c_1^2\sigma \geqslant a_0 \vee \varepsilon^{2/3}$ and $h > 2\sigma$, then*

$$\begin{aligned}
\mathbb{P}^{-c_1\sqrt{\sigma},x_0}&\bigl\{0 < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)}\ \ \forall s \in [-c_1\sqrt{\sigma}, t_1]\bigr\} \\
&\leqslant 2\exp\biggl\{-\bar{\kappa}\,\frac{1}{\log(h/\sigma)}\frac{\alpha(t_1,-c_1\sqrt{\sigma})}{\varepsilon}\biggr\}
\end{aligned} \tag{3.96}$$

holds with $\alpha(t_1,-c_1\sqrt{\sigma}) = \int_{-c_1\sqrt{\sigma}}^{t} a(s)\,\mathrm{d}s$, for $-c_1\sqrt{\sigma} \leqslant t_1 \leqslant c_1\sqrt{\sigma}$ and all initial conditions x_0 satisfying $0 \leqslant x_0 \leqslant x^{\det}_{-c_1\sqrt{\sigma}} + h\sqrt{\zeta(-c_1\sqrt{\sigma})}$.

Proof:

1. Let $\varrho = \varrho(h/\sigma) \geqslant 1$ and define a partition $-c_1\sqrt{\sigma} = u_0 < \dots < u_K = t_1$ of $[-c_1\sqrt{\sigma}, t_1]$ by

$$\alpha(u_k, u_{k-1}) = \varrho\varepsilon \qquad \text{for } 1 \leqslant k < K = \left\lceil \frac{\alpha(t_1, -c_1\sqrt{\sigma})}{\varrho\varepsilon} \right\rceil. \tag{3.97}$$

We would like to control the probability of not reaching the t-axis during the time interval $[u_k, u_{k+1}]$. Let

$$Q_k = \sup_{0<x_k\leqslant x^{\det}_{u_k}+h\sqrt{\zeta(u_k)}} \mathbb{P}^{u_k,x_k}\bigl\{0 < x_s \leqslant x^{\det}_s + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, u_{k+1}]\bigr\}. \tag{3.98}$$

Then the probability on the left-hand side of (3.96) is

$$\begin{aligned}
&\mathbb{P}^{-c_1\sqrt{\sigma},x_0}\bigl\{0 < x_s \leqslant x^{\det}_s + h\sqrt{\zeta(s)} \;\; \forall s \in [-c_1\sqrt{\sigma}, t_1]\bigr\} \\
&\quad = \mathbb{E}^{-c_1\sqrt{\sigma},x_0}\Bigl\{1_{\{0<x_s\leqslant x^{\det}_s+h\sqrt{\zeta(s)} \;\; \forall s\in[-c_1\sqrt{\sigma},u_{K-1}]\}} \\
&\qquad\qquad \mathbb{P}^{u_{K-1},x_{u_{K-1}}}\bigl\{0 < x_s \leqslant x^{\det}_s + h\sqrt{\zeta(s)} \;\; \forall s \in [u_{K-1}, u_K]\bigr\}\Bigr\} \\
&\quad \leqslant Q_{K-1}\, \mathbb{P}^{-c_1\sqrt{\sigma},x_0}\bigl\{0 < x_s \leqslant x^{\det}_s + h\sqrt{\zeta(s)} \;\; \forall s \in [-c_1\sqrt{\sigma}, u_{K-1}]\bigr\} \\
&\quad \leqslant \dots \leqslant \prod_{k=0}^{K-1} Q_k.
\end{aligned} \tag{3.99}$$

If we manage to estimate each Q_k by a constant less than 1 (say, 1/2), then the probability will be exponentially small in K. In the sequel, we shall estimate Q_k uniformly in $k = 0, \dots K-2$, and bound Q_{K-1} by 1, since the last interval of the partition may be too small to get a good bound. So let $k < K-1$ from now on.

2. We consider first the case $0 < x_k \leqslant x^{\det}_{u_k}$. We define the process $(x^{(k)}_s)_{u_k\leqslant s\leqslant u_{k+1}}$ as the solution of the linearized SDE

$$\mathrm{d}x^{(k)}_s = a(s)x^{(k)}_s\,\mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W^{(k)}_s, \qquad x^{(k)}_{u_k} = x_k, \tag{3.100}$$

where $(W^{(k)}_s)_{s\in[u_k,u_{k+1}]}$ is the Brownian motion $W^{(k)}_s = W_s - W_{u_k}$. Let $v^{(k)}_{u_{k+1}}$ denote the variance of $x^{(k)}_{u_{k+1}}$. Then

$$\begin{aligned}
\mathrm{e}^{-2\alpha(u_{k+1},u_k)/\varepsilon}\, v^{(k)}_{u_{k+1}} &= \frac{\sigma^2}{\varepsilon}\int_{u_k}^{u_{k+1}} \mathrm{e}^{-2\alpha(u,u_k)/\varepsilon}\,\mathrm{d}u \\
&\geqslant \frac{\sigma^2}{2} \inf_{u_k\leqslant u\leqslant u_{k+1}} \frac{1}{a(u)}\bigl[1 - \mathrm{e}^{-2\alpha(u_{k+1},u_k)/\varepsilon}\bigr] \\
&\geqslant \frac{1-\mathrm{e}^{-2\varrho}}{2}\,\frac{\sigma^2}{a(u_k)\vee a(u_{k+1})}.
\end{aligned} \tag{3.101}$$

We can now apply Lemma 3.11 in the particular case $x^{\det}_s \equiv 0$ to show that if $0 < x_s \leqslant \delta$ for $s \in [u_k, u_{k+1}]$, then $x^{(k)}_s \geqslant x_s$ in the same interval. We thus obtain

$$\mathbb{P}^{u_k,x_k}\bigl\{0 < x_s \leqslant x^{\det}_s + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, u_{k+1}]\bigr\} \leqslant \mathbb{P}^{u_k,x_k}\bigl\{x^{(k)}_s > 0 \;\; \forall s \in [u_k, u_{k+1}]\bigr\}. \tag{3.102}$$

The probability on the right-hand side satisfies

$$\begin{aligned}\mathbb{P}^{u_k,x_k}\bigl\{x_s^{(k)} > 0 \ \ \forall s \in [u_k, u_{k+1}]\bigr\} &= 1 - 2\mathbb{P}^{u_k,x_k}\bigl\{x_{u_{k+1}}^{(k)} < 0\bigr\} \\ &= 2\mathbb{P}^{u_k,x_k}\bigl\{x_{u_{k+1}}^{(k)} \geqslant 0\bigr\} - 1,\end{aligned} \tag{3.103}$$

yielding

$$\begin{aligned}&\mathbb{P}^{u_k,x_k}\bigl\{0 < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \ \ \forall s \in [u_k, u_{k+1}]\bigr\} \\ &\leqslant \frac{2}{\sqrt{2\pi}} \int_0^{x_k \,\mathrm{e}^{\alpha(u_{k+1},u_k)/\varepsilon}\left(v_{u_{k+1}}^{(k)}\right)^{-1/2}} \mathrm{e}^{-z^2/2}\,\mathrm{d}z \\ &\leqslant \frac{2}{\sqrt{2\pi}} \frac{x_k}{\sqrt{\mathrm{e}^{-2\alpha(u_{k+1},u_k)/\varepsilon}\, v_{u_{k+1}}^{(k)}}} \\ &\leqslant \frac{2}{\sqrt{\pi}} \sqrt{\frac{1}{1-\mathrm{e}^{-2\varrho}}} \frac{\sqrt{a(u_k) \vee a(u_{k+1})}}{\sigma} x_{u_k}^{\det}.\end{aligned} \tag{3.104}$$

By making c_1 small enough, we can guarantee that this bound is smaller than some imposed constant of order 1, say 1/2. This shows that the length of $[u_k, u_{k+1}]$ has been chosen large enough that the probability of reaching the t-axis during this time interval is appreciable.

3. We examine now the case $x_{u_k}^{\det} < x_k < x_{u_k}^{\det} + h\sqrt{\zeta(s)}$. We introduce a time $\tilde{u}_k \in (u_k, u_{k+1})$, defined by

$$\alpha(\tilde{u}_k, u_k) = \frac{1}{2}\varrho\varepsilon. \tag{3.105}$$

Our strategy will be to show that x_t is likely to cross $x_t^{\det}$ before time $\tilde{u}_k$, which will allow us to use the previous result. Proposition 3.7 implies the existence of a constant $L > 0$ such that

$$\frac{1}{2}\varrho\varepsilon = \int_{u_k}^{\tilde{u}_k} a(u)\,\mathrm{d}u \leqslant L \int_{u_k}^{\tilde{u}_k} (-\bar{a}(u))\,\mathrm{d}u = L|\bar{\alpha}(\tilde{u}_k, u_k)|. \tag{3.106}$$

Let $(y_s^{(k)})_{u_k \leqslant s \leqslant u_{k+1}}$ be the solution of the linear SDE

$$\mathrm{d}y_s^{(k)} = \bar{a}(s) y_s^{(k)}\,\mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_s^{(k)}, \qquad y_{u_k}^{(k)} = y_k = x_k - x_{u_k}^{\det}, \tag{3.107}$$

where $(W_s^{(k)})_{s\in[u_k,u_{k+1}]}$ is again the Brownian motion $W_s^{(k)} = W_s - W_{u_k}$. The variance of $y_{\tilde{u}_k}^{(k)}$ is

$$\tilde{v}_{\tilde{u}_k}^{(k)} = \frac{\sigma^2}{\varepsilon} \int_{u_k}^{\tilde{u}_k} \mathrm{e}^{2\bar{\alpha}(\tilde{u}_k,s)/\varepsilon}\,\mathrm{d}s \geqslant \frac{1-\mathrm{e}^{-\varrho/L}}{2} \frac{\sigma^2}{|\bar{a}(u_k)| \vee |\bar{a}(\tilde{u}_k)|}. \tag{3.108}$$

Lemma 3.11 shows that if $x_s^{\det} \leqslant x_s \leqslant \delta$ on the interval $[u_k, \tilde{u}_k]$, then $x_s - x_s^{\det} \leqslant y_s^{(k)}$ on that interval. If we introduce the stopping time

$$\tau_k = \inf\bigl\{s \in [u_k, \tilde{u}_k] \colon x_s = x_s^{\det}\bigr\} \in [u_k, u_{k+1}] \cup \{\infty\}, \tag{3.109}$$

then we have

$$
\begin{aligned}
&\mathbb{P}^{u_k,x_k}\bigl\{0 < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, u_{k+1}]\bigr\} \\
&\quad \leqslant \mathbb{P}^{u_k,x_k}\bigl\{x_s^{\det} < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, \tilde u_k]\bigr\} \\
&\qquad + \mathbb{E}^{u_k,x_k}\Bigl\{1_{\{\tau_k < \tilde u_k\}} \mathbb{P}^{\tau_k, x_{\tau_k}^{\det}}\bigl\{0 < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [\tau_k, u_{k+1}]\bigr\}\Bigr\}. \qquad (3.110)
\end{aligned}
$$

The second term on the right-hand side can be bounded, as in (3.104), by

$$
\frac{2}{\sqrt{\pi}} \sqrt{\frac{1}{1 - \mathrm{e}^{-\varrho}}} \frac{\sqrt{a(u_k) \vee a(u_{k+1})}}{\sigma} \sup_{u_k \leqslant s \leqslant \tilde u_k} x_s^{\det}. \qquad (3.111)
$$

Using (3.108), the first term on the right-hand side of (3.110) can be estimated in the following way:

$$
\begin{aligned}
&\mathbb{P}^{u_k,x_k}\bigl\{x_s^{\det} < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, \tilde u_k]\bigr\} \\
&\quad \leqslant \mathbb{P}^{u_k,y_k}\bigl\{y_s^{(k)} > 0 \;\; \forall s \in [u_k, \tilde u_k]\bigr\} \\
&\quad \leqslant \frac{2}{\sqrt{2\pi}} \frac{y_k \,\mathrm{e}^{\overline{\alpha}(\tilde u_k, u_k)/\varepsilon}}{\bigl(\tilde v_{\tilde u_k}^{(k)}\bigr)^{1/2}} \\
&\quad \leqslant \frac{2}{\sqrt{\pi}} \sqrt{\frac{1}{1 - \mathrm{e}^{-\varrho/L}}} \sqrt{|\bar a(u_k)| \vee |\bar a(\tilde u_k)|} \,\frac{h}{\sigma} \sqrt{\zeta(u_k)} \,\mathrm{e}^{-\varrho/2L}. \qquad (3.112)
\end{aligned}
$$

Using Proposition 3.7 and (3.46), it is easy to show that the expression $(|\bar a(u_k)| \vee |\bar a(\tilde u_k)|)\zeta(u_k)$ is uniformly bounded by a constant independent of k and ε. The sum of (3.111) and of the last term in (3.112) provides an upper bound for Q_k.

4. Using the fact that for $|u| \leqslant c_1\sqrt{\sigma}$ and $c_1^2\sigma \geqslant a_0 \vee \varepsilon^{2/3}$, one has $a(u) = \mathcal{O}(c_1^2\sigma)$ and $x_u^{\det} = \mathcal{O}(c_1\sqrt{\sigma})$, we arrive at the bound

$$
Q_k \leqslant C\Bigl(c_1^2 + \frac{h}{\sigma}\,\mathrm{e}^{-\varrho/2L}\Bigr), \qquad (3.113)
$$

where the constant C can be chosen independent of ϱ because $\varrho \geqslant 1$ by assumption. Thus if we choose $c_1^2 \leqslant 1/4C$ and $\varrho = 2L\log(4Ch/\sigma) \vee 1$, we obtain that $Q_k \leqslant 1/2$ for $k = 0, \dots, K-2$. This yields

$$
\prod_{k=0}^{K-1} Q_k \leqslant 2\frac{1}{2^K} \leqslant 2\exp\Bigl\{-(\log 2)\,\frac{\alpha(t_1, -c_1\sqrt{\sigma})}{\varrho\varepsilon}\Bigr\}, \qquad (3.114)
$$

and the result follows from our choice of ϱ. □

Now the proof of Theorem 2.3 follows from (3.76) and the two preceding propositions, where we use the Markov property to "restart" at time $-c_1\sqrt{\sigma}$ before applying Proposition 3.13.

3.4 Escape from the saddle

In this subsection, we investigate the behaviour of the random motion x_t given by the SDE (3.1) for $t \geqslant t_2 \geqslant c_1\sqrt{\sigma}$, i.e., after the transition regime. We want to show that x_t is likely to leave a suitably defined neighbourhood of the saddle within time $\mathcal{O}(\varepsilon|\log\sigma|/t_2^2)$.

The proof of Theorem 2.4 is very similar to the proof of [BG, Theorem 2.9], and for the sake of brevity, we will refrain from giving all the details. Instead, we will discuss how to proceed and then focus on those parts which need to be modified.

From now on, we will assume that $t \geqslant t_2 \geqslant c_1\sqrt{\sigma}$ and that σ is large enough in order to allow for transitions, i. e., $\sigma \geqslant a_0 \vee \varepsilon^{2/3}$. We want to estimate the first exit time $\tau_{\mathcal{D}(\kappa)}$ of x_t from the set

$$\mathcal{D}(\kappa) = \Big\{(x,t) \in [-\delta,\delta] \times [c_1\sqrt{\sigma}, T] \colon \frac{f(x,t)}{x} > \kappa a(t)\Big\}, \tag{3.115}$$

where $\kappa \in (0,1)$ is a constant. Note that the upper boundary $\tilde{x}(t)$ of $\mathcal{D}(\kappa)$ satisfies $\tilde{x}(t) = \sqrt{1-\kappa}(1-\mathcal{O}(t))x^\star(t)$. Our first step towards estimating $\tau_{\mathcal{D}(\kappa)}$ is to estimate the first exit time $\tau_{\mathcal{S}}$ from a smaller strip $\mathcal{S}$, defined by

$$\mathcal{S} = \Big\{(x,t) \in [-\delta,\delta] \times [c_1\sqrt{\sigma}, T] \colon |x| < \frac{h}{\sqrt{a(s)}}\Big\}, \tag{3.116}$$

where we will choose h later. Note that $h <$ *const* σ for some (small) constant would assure $\mathcal{S} \subset \mathcal{D}(\kappa)$. We will not impose such a restrictive condition on h but replace $\mathcal{S}$ by $\mathcal{S} \cap \mathcal{D}(\kappa)$ in case $\mathcal{S}$ is not a subset of $\mathcal{D}(\kappa)$. The following proposition gives our estimate on the first exit time from $\mathcal{S}$.

Proposition 3.14. *Let $t_2 \geqslant c_1\sqrt{\sigma}$ and $(x_2, t_2) \in \mathcal{S}$. Then there exists a constant $L > 0$ such that for any $\mu > 0$, we have*

$$\mathbb{P}^{t_2,x_2}\big\{\tau_{\mathcal{S}} \geqslant t\big\} \leqslant \Big(\frac{h}{\sigma}\Big)^\mu \exp\Big\{-\frac{\mu}{1+\mu}\frac{\alpha(t,t_2)}{\varepsilon}\Big[1-\mathcal{O}\Big(\frac{1}{\mu\log(h/\sigma)}\Big)\Big]\Big\} \tag{3.117}$$

under the condition

$$\Big(\frac{h}{\sigma}\Big)^{3+\mu}\Big(1+(1+\mu)\frac{\varepsilon}{t_2^3}\log\frac{h}{\sigma}\Big) \leqslant L\frac{t_2^4}{\sigma^2}. \tag{3.118}$$

Proof: The proof follows along the lines of the one of [BG, Proposition 4.7], the main difference being the quadratic behaviour $a_- t^2 \leqslant a(t) \leqslant a_+ t^2$ of a in our case as opposed to the linear one in [BG].

We start by defining a partition $t_2 = u_0 < \cdots < u_K = t$ of $[t_2, t]$, given by

$$\alpha(u_k, u_{k-1}) = (1+\mu)\frac{\varepsilon}{2}\log\frac{h^2}{\sigma^2}, \qquad \text{for } 1 \leqslant k < K = \Big\lceil \frac{2\alpha(t,t_2)}{(1+\mu)\varepsilon\log(h^2/\sigma^2)}\Big\rceil. \tag{3.119}$$

On each interval $[u_k, u_{k+1}]$, we consider a Gaussian approximation $(x_t^{(k)})_{t\in[u_k,u_{k+1}]}$ of x_t, defined by

$$\mathrm{d}x_t^{(k)} = \frac{1}{\varepsilon}a(t)x_t^{(k)}\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t^{(k)} \qquad x_{u_k}^{(k)} = x_{u_k}, \tag{3.120}$$

where $W_t^{(k)} = W_t - W_{u_k}$. If $|x_s|\sqrt{a(s)} \leqslant h$ for all $s \in [u_k, u_{k+1}]$, then by (3.2) and (3.3), there is a constant $M > 0$ such that

$$\begin{aligned} |x_s - x_s^{(k)}| &\leqslant \frac{1}{\varepsilon}\int_{u_k}^s |g_0(x_u,u)x_u|\,\mathrm{e}^{\alpha(s,u)/\varepsilon}\,\mathrm{d}u \\ &\leqslant M\frac{h^3}{a(u_k)^{3/2}}\frac{1}{a(u_k)}\,\mathrm{e}^{\alpha(u_{k+1},u_k)/\varepsilon} \leqslant \frac{h}{\sqrt{a(s)}} \end{aligned} \tag{3.121}$$

for all $s \in [u_k, u_{k+1}]$, provided the condition

$$h^2 \leqslant \frac{a_-^2}{M} \sqrt{\frac{a(u_k)}{a(u_{k+1})}} \, \mathrm{e}^{-\alpha(u_{k+1},u_k)/\varepsilon} \, t_2^4 \tag{3.122}$$

holds for all k. Now,

$$\frac{a(u_{k+1})}{a(u_k)} \leqslant 1 + \frac{ca_+}{a_-^3} \frac{\alpha(u_{k+1},u_k)}{t_2^3} \Big(1 + \frac{a_+}{a_-^2} \frac{\alpha(u_{k+1},u_k)}{t_2^3} \Big), \tag{3.123}$$

where c is a constant satisfying $0 \leqslant a'(t) \leqslant ct$ for all $t \in [0,T]$. This shows that there exists a constant $L > 0$ such that the condition (3.122) is satisfied whenever

$$\Big(\frac{h}{\sigma}\Big)^{3+\mu} \Big(1 + \frac{\alpha(u_{k+1},u_k)}{t_2^3} \Big) \leqslant L \frac{t_2^4}{\sigma^2}. \tag{3.124}$$

This condition is equivalent to (3.118).

Assume $|x_{u_k}|\sqrt{a(u_k)} \leqslant h$ for the moment. Then,

$$\begin{aligned}
\mathbb{P}^{u_k,x_{u_k}} \Big\{ \sup_{u_k \leqslant s \leqslant u_{k+1}} |x_s| \sqrt{a(s)} \leqslant h \Big\} &\leqslant \mathbb{P}^{u_k,x_{u_k}} \Big\{ |x^{(k)}_{u_{k+1}}| \sqrt{a(u_{k+1})} \leqslant 2h \Big\} \\
&\leqslant \frac{4h}{\sqrt{2\pi v^{(k)}_{u_{k+1}} a(u_{k+1})}},
\end{aligned} \tag{3.125}$$

where $v^{(k)}_{u_{k+1}}$ denotes the variance of $x^{(k)}_{u_{k+1}}$. By partial integration, we find

$$v^{(k)}_{u_{k+1}} = \frac{\sigma^2}{\varepsilon} \int_{u_k}^{u_{k+1}} \mathrm{e}^{2\alpha(u_{k+1},s)/\varepsilon} \, \mathrm{d}s \geqslant \frac{\sigma^2}{a(u_{k+1})} \Big[\mathrm{e}^{2\alpha(u_{k+1},u_k)/\varepsilon} - 1 \Big]. \tag{3.126}$$

Now, the Markov property yields

$$\mathbb{P}^{t_2,x_2} \big\{ \tau_{\mathcal{S}} \geqslant t \big\} = \mathbb{P}^{t_2,x_2} \Big\{ \sup_{t_2 \leqslant s \leqslant t} |x_s| \sqrt{a(s)} \leqslant h \Big\} \leqslant \prod_{k=0}^{K-1} \bigg(\frac{4}{\sqrt{2\pi}} \frac{h}{\sqrt{v^{(k)}_{u_{k+1}} a(u_{k+1})}} \wedge 1 \bigg), \tag{3.127}$$

and the bound (3.117) follows by a straightforward calculation. □

The preceding proposition shows that a path starting in $\mathcal{S}$ is likely to leave $\mathcal{S}$ after a short time. We want to show that such a path (or any path starting in $\mathcal{D}(\kappa) \setminus \mathcal{S}$) is also likely to leave $\mathcal{D}(\kappa)$. For this purpose, we will again compare x_t to a Gaussian approximation, given by

$$\mathrm{d}x_t^0 = \frac{1}{\varepsilon} a_0(t) x_t^0 \, \mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}} \, \mathrm{d}W_t, \tag{3.128}$$

where $a_0(t) = \kappa a(t)$, so that $f(x,t)/x \geqslant a_0(t)$ in $\mathcal{D}(\kappa)$. Assume that $x_{t_2} > 0$. Then $x_s \geqslant x_s^0$ holds as long as x_s neither leaves $\mathcal{D}(\kappa)$ nor crosses the t-axis, cf. [BG, Lemma 4.8]. Therefore we can proceed as follows. Once a path is in $\mathcal{D}(\kappa) \setminus \mathcal{S}$, there are two possibilities. Either, x_s^0 does not return to zero, or it does. If x_t^0 does not return to zero, then it is likely to leave $\mathcal{D}(\kappa)$ via the upper boundary and so is x_t. So we are left with the case of

x_t^0 returning to zero. This event has a small but not negligible probability. Note that if x_t^0 returns to zero, then x_t is still non-negative. If x_t has nevertheless left $\mathcal{D}(\kappa)$, we are done. If not, x_t is either in $\mathcal{S}$ or in $\mathcal{D}(\kappa) \setminus \mathcal{S}$. Since we may assume that, after a short time, x_t is in $\mathcal{D}(\kappa) \setminus \mathcal{S}$ again, we can repeat the above argument.

Making the above-said precise, we obtain an integral equation for an upper bound on the probability that x_u does not leave $\mathcal{D}(\kappa)$ up to time t, which will be solved by iterations. We will cite the integral equation from [BG], as the general arguments leading to it do not require adaptation. Let us first introduce the necessary notations. We choose $h = K\sigma$ for some (possibly large) constant $K > 0$. For $\kappa \in (0,1)$, we choose $\mu > 0$ in such a way that

$$\frac{1}{2}\frac{\mu}{1+\mu}\Big[1 - \mathcal{O}\Big(\frac{1}{\mu \log K}\Big)\Big] \leqslant \kappa \leqslant \frac{\mu}{1+\mu}\Big[1 - \mathcal{O}\Big(\frac{1}{\mu \log K}\Big)\Big] \tag{3.129}$$

for all large enough K. Note that choosing κ too close to 1 requires large μ and is thus not desirable. Since we want to apply Proposition 3.14 on the first exit time from $\mathcal{S}$ with $t_2 \geqslant c_2\sqrt{\sigma}$ for a suitably chosen c_2, Condition (3.118) must be satisfied. Therefore, we choose $c_2 = c_2(K)$ large enough for

$$K^{3+\mu}\Big(1 + \frac{1+\mu}{c_2^3}\log K\Big) \leqslant Lc_2^4 \tag{3.130}$$

to hold. Now, set

$$g(t,s) = \frac{\mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon}}{\sqrt{1 - \mathrm{e}^{-2\kappa\alpha(t,s)/\varepsilon}}}, \tag{3.131}$$

and

$$C = \max\Big\{\frac{\tilde{x}(t)\sqrt{\kappa a(t)}}{\sqrt{\pi}\sigma}, 1\Big\} \qquad \text{and} \qquad c = \sqrt{\pi\kappa}\Big(\frac{h}{\sigma}\Big)^{1+\mu}\mathrm{e}^{-\kappa h^2/\sigma^2}. \tag{3.132}$$

For $t_2 \leqslant s \leqslant t \leqslant T$, let $Q_t^{(0)}(s) \equiv 1$, and define $Q_t^{(n)}(s)$ for $n \geqslant 1$ by

$$Q_t^{(n)}(s) = Cg(t,s) + c\,\mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} + c\int_s^t Q_t^{(n-1)}(u)\frac{a(u)}{\varepsilon}\,\mathrm{e}^{-\kappa\alpha(u,s)/\varepsilon}\big[1 + g(u,s)\big]\,\mathrm{d}u. \tag{3.133}$$

Then, [BG, (4.95) and (4.107)] show that for any n and $s \geqslant t_2 \geqslant c_2\sqrt{\sigma}$,

$$\sup_{x\colon (x,s)\in\mathcal{S}} \mathbb{P}^{s,x}\big\{\tau_{\mathcal{D}(\kappa)} \geqslant t\big\} \leqslant 2\Big(\frac{h}{\sigma}\Big)^{\mu}\mathrm{e}^{-\kappa\alpha(t,s)/\varepsilon} + \kappa\Big(\frac{h}{\sigma}\Big)^{\mu}\int_s^t Q_t^{(n)}(u)\frac{a(u)}{\varepsilon}\,\mathrm{e}^{-\kappa\alpha(u,s)/\varepsilon}\,\mathrm{d}u \tag{3.134}$$

and

$$\sup_{x\colon (x,s)\in\mathcal{D}(\kappa)\setminus\mathcal{S}} \mathbb{P}^{s,x}\big\{\tau_{\mathcal{D}(\kappa)} \geqslant t\big\} \leqslant Q_t^{(n)}(s). \tag{3.135}$$

Next we estimate $Q_t^{(n)}$ by showing that

$$Q_t^{(n)}(s) \leqslant Cg(t,s) + a_n\,\mathrm{e}^{-\kappa\alpha(t,s)/2\varepsilon} + b_n \qquad \forall n \tag{3.136}$$

holds with $a_1 = c$, $b_1 = 3c/\kappa$ in the case $n = 1$, and with

$$a_n = c\Big(1 + \frac{4C}{\kappa}\Big)\sum_{j=0}^{n-2}\Big(\frac{6c}{\kappa}\Big)^j + c\Big(\frac{6c}{\kappa}\Big)^{n-1} \leqslant \Big(1 + \frac{4C}{\kappa}\Big)\frac{c}{1 - 6c/\kappa} \leqslant 2c\Big(1 + \frac{4C}{\kappa}\Big) \tag{3.137}$$

$$b_n = \Big(\frac{3c}{\kappa}\Big)^n \tag{3.138}$$

for $n > 1$, provided $6c/\kappa \leqslant 1/2$. Note that the latter imposes a condition on $K = h/\sigma$. To obtain the bound (3.136), we proceed as in [BG], the only difference lying in the term $a_n\,\mathrm{e}^{-\kappa\alpha(t,s)/2\varepsilon}$, where we sacrifice a factor of 2 in the exponent in order to gain a smaller coefficient a_n. Our choice of a_n yields a less restrictive condition on h/σ, namely we only need

$$\frac{12\sqrt{\pi}}{\sqrt{\kappa}}\Bigl(\frac{h}{\sigma}\Bigr)^{1+\mu}\,\mathrm{e}^{-\kappa h^2/\sigma^2} \leqslant 1, \tag{3.139}$$

which is satisfied whenever $K = h/\sigma$ is large enough.

Now, (3.137) and (3.138) imply that for K and $c_2(K)$ large enough,

$$\sup_{x\colon (x,t_2)\in\mathcal{D}(\kappa)} \mathbb{P}^{t_2,x}\bigl\{\tau_{\mathcal{D}(\kappa)} \geqslant t\bigr\} \leqslant C_0\Bigl(\frac{t}{\sqrt{\sigma}}\Bigr)^2 \frac{\mathrm{e}^{-\kappa\alpha(t,t_2)/2\varepsilon}}{\sqrt{1-\mathrm{e}^{-\kappa\alpha(t,t_2)/\varepsilon}}} \qquad \text{for all } t \geqslant t_2 \geqslant c_2\sqrt{\sigma} \tag{3.140}$$

with some constant C_0. This completes our outline of the proof of Theorem 2.4.

4 Asymmetric case

We consider in this section the nonlinear SDE

$$\mathrm{d}x_t = \frac{1}{\varepsilon} f(x_t,t)\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \tag{4.1}$$

where f satisfies the hypotheses given at the beginning of Subsection 2.3. By rescaling x, we can arrange for $\partial_{xx} f(x_{\mathrm{c}},0) = -2$, so that Taylor's formula allows us to write

$$\partial_x f(x_{\mathrm{c}}+\tilde{x},t) = \partial_x f(x_{\mathrm{c}},t) + \tilde{x}\bigl[-2 + r_1(\tilde{x},t)\bigr] \tag{4.2}$$

where $r_1 \in \mathcal{C}^1$ and $r_1(0,0) = 0$. Since $\partial_x f(x_{\mathrm{c}},t) = \mathcal{O}(t^2)$ by assumption, $\partial_x f(x,t)$ vanishes on a curve $\bar{x}(t) = x_{\mathrm{c}} + \mathcal{O}(t^2)$. We further obtain that

$$\begin{aligned} f(\bar{x}(t)+z,t) &= f(\bar{x}(t),t) + z^2\bigl[-1 + r_0(z,t)\bigr] \\ f(\bar{x}(t),t) &= f(x_{\mathrm{c}},t) + \mathcal{O}(t^4) = a_0 + a_1 t^2 + \mathcal{O}(t^3), \end{aligned} \tag{4.3}$$

where $r_0 \in \mathcal{C}^1$ and $r_0(0,0) = 0$. Thus $f(x,t)$ vanishes on two curves $x^\star_+(t)$ and $x^\star_0(t)$, which behave near $t = 0$ like $x_{\mathrm{c}} \pm \sqrt{a_0 + a_1 t^2}[1 + \mathcal{O}(\sqrt{a_0 + a_1 t^2})]$, as indicated in (2.34). The behaviour of the linearization follows from (4.2).

4.1 Deterministic case

The proof of Theorem 2.5 follows closely the proof of Theorem 2.1, with some minor differences we comment on here. The dynamics of $y_t = x_t - x^\star_+(t)$ is still governed by an equation of the form

$$\varepsilon\frac{\mathrm{d}y}{\mathrm{d}t} = a^\star_+(t)y + b^\star_+(y,t) - \varepsilon\frac{\mathrm{d}x^\star_+}{\mathrm{d}t}, \tag{4.4}$$

but now Taylor's formula yields the relations

$$a^\star_+(t) \asymp -(\sqrt{a_0} \vee |t|) \tag{4.5}$$

$$b^\star_+(y,t) = -y^2\bigl[1 + \mathcal{O}(\sqrt{a_0}) + \mathcal{O}(t) + \mathcal{O}(y)\bigr], \tag{4.6}$$

while relation (3.10) holds for the derivative of $x^\star_+$, with $t^\star$ replaced by $t^\star_+$. Lemma 3.2 becomes

Lemma 4.1. *Let $\tilde{a}(t)$ be a continuous function satisfying $\tilde{a}(t) \asymp -(\beta \vee |t|)$ for $|t| \leqslant T$, where $\beta = \beta(\varepsilon) \geqslant 0$. Let $\chi_0 \asymp 1$, and define $\tilde{\alpha}(t,s) = \int_s^t \tilde{a}(u)\,\mathrm{d}u$. Then*

$$\chi_0\,\mathrm{e}^{\tilde{\alpha}(t,-T)/\varepsilon} + \frac{1}{\varepsilon}\int_{-T}^{t} \mathrm{e}^{\tilde{\alpha}(t,s)/\varepsilon}\,\mathrm{d}s \asymp \begin{cases} \dfrac{1}{\beta \vee \sqrt{\varepsilon}} & \text{for } |t| \leqslant \beta \vee \sqrt{\varepsilon} \\ \dfrac{1}{|t|} & \text{for } \beta \vee \sqrt{\varepsilon} \leqslant |t| \leqslant T. \end{cases} \tag{4.7}$$

Proposition 3.3 carries over with some obvious adjustments, and shows the existence of a constant c_0 such that

$$y_t \asymp \frac{\varepsilon}{|t|} \qquad \text{for } -T \leqslant t \leqslant t_0 = -c_0(\sqrt{a_0} \vee \sqrt{\varepsilon}). \tag{4.8}$$

In particular, $y_{t_0} \asymp (\varepsilon/\sqrt{a_0}) \wedge \sqrt{\varepsilon}$. An adaptation of Proposition 3.4 yields the existence of a constant $\gamma_0 > 0$ such that, for $a_0 \geqslant \gamma_0\varepsilon$,

$$y_t = C_1(t)(t_+^\star - t) + C_2(t) \qquad \text{with} \qquad C_1(t) \asymp \frac{\varepsilon}{a_0}, \quad C_2(t) \asymp \frac{\varepsilon^2}{a_0^{3/2}} \tag{4.9}$$

for all $|t| \leqslant |t_0|$. This shows in particular that y_t vanishes at a time $\tilde{t}$ satisfying $\tilde{t} - t_+^\star \asymp \varepsilon/\sqrt{a_0}$. Proposition 3.5 is replaced by

Proposition 4.2. *Assume that $a_0 < \gamma_0\varepsilon$. Then, for any fixed $t_1 \asymp \sqrt{\varepsilon}$,*

$$x_t - x_{\mathrm{c}} \asymp \sqrt{\varepsilon} \qquad \text{for } t_0 \leqslant t \leqslant t_1, \tag{4.10}$$

and x_t crosses $x_+^\star(t)$ at a time $\tilde{t}$ satisfying $\tilde{t} \asymp \sqrt{\varepsilon}$.

PROOF: Let $\tilde{x}_t = x_t - x_{\mathrm{c}}$. We first observe that by Taylor's formula,

$$f(x_{\mathrm{c}} + \tilde{x}, t) = f(x_{\mathrm{c}}, t) + \tilde{x}\partial_x f(x_{\mathrm{c}}, t) + \tilde{x}^2\bigl[-1 + \mathcal{O}(\tilde{x}) + \mathcal{O}(t)\bigr]. \tag{4.11}$$

This shows that

$$\varepsilon\frac{\mathrm{d}\tilde{x}}{\mathrm{d}t} = a_0 + a_1 t^2 - \tilde{x}^2 + \mathcal{O}(t^3) + \mathcal{O}(t^2\tilde{x}) + \mathcal{O}(t\tilde{x}^2) + \mathcal{O}(\tilde{x}^3). \tag{4.12}$$

Thus, with the rescaling

$$\tilde{x} = a_1^{1/4}\sqrt{\varepsilon}\,z, \qquad t = a_1^{-1/4}\sqrt{\varepsilon}\,s, \tag{4.13}$$

we obtain that z_t obeys a perturbation of order $\sqrt{\varepsilon}$ of the Riccati equation

$$\frac{\mathrm{d}z}{\mathrm{d}s} = \tilde{a}_0 + s^2 - z^2, \qquad \text{with } \tilde{a}_0 = \frac{1}{\sqrt{a_1}}\frac{a_0}{\varepsilon} < \frac{\gamma_0}{\sqrt{a_1}}. \tag{4.14}$$

One easily shows that the solution satisfies $z_s \asymp 1$ for s of order 1, and this property carries over to the perturbed equation with the help of Gronwall's inequality. Finally, since $\tilde{x}_t \asymp \sqrt{\varepsilon}$ and $x_+^\star(t) - x_{\mathrm{c}} \asymp \sqrt{a_0} \vee |t|$, these curves necessarily cross at a time $\tilde{t} \asymp \sqrt{\varepsilon}$. $\square$

The assertion on the existence of a particular solution $\widehat{x}^{\text{det}}$ tracking the unstable equilibrium branch $x_0^\star(t)$ follows from the observation that $z_s = x_{-s}$ satisfies the equation

$$\varepsilon \frac{\mathrm{d}z_s}{\mathrm{d}s} = -f(z_s, -s). \tag{4.15}$$

This system admits $z_0^\star(s) = x_0^\star(-s)$ as a stable equilibrium branch. Thus the same arguments as above can be used to show the existence of a solution z_s tracking $z_0^\star(s)$, with similar properties. Proposition 3.7 admits the following counterpart:

Proposition 4.3. *For all $t \in [-T, T]$ and all $a_0 = \mathcal{O}_\varepsilon(1)$,*

$$\bar{a}(t) := \partial_x f(x_t^{\text{det}}, t) \asymp -(|t| \vee \sqrt{a_0} \vee \sqrt{\varepsilon}) \tag{4.16}$$

$$\widehat{a}(t) := \partial_x f(\widehat{x}_t^{\text{det}}, t) \asymp |t| \vee \sqrt{a_0} \vee \sqrt{\varepsilon}. \tag{4.17}$$

PROOF: The proof is a direct consequence of (4.2) and the properties of x_t^{det}, and thus much simpler than the proof of Proposition 3.7. □

Finally, with Lemma 4.1, we immediately obtain

$$\zeta(t) := \frac{1}{2|\bar{a}(-T)|} \,\mathrm{e}^{2\overline{\alpha}(t,-T)/\varepsilon} + \frac{1}{\varepsilon} \int_{-T}^{t} \mathrm{e}^{2\overline{\alpha}(t,s)/\varepsilon} \,\mathrm{d}s \asymp \frac{1}{|t| \vee \sqrt{a_0} \vee \sqrt{\varepsilon}}. \tag{4.18}$$

4.2 The transition regime

We consider now the regime of σ sufficiently large to allow for transitions from the potential well at $x_+^\star$ to the potential well at $x_-^\star$, by passing over the saddle at $x_0^\star$. Here x_t^{det} and $\widehat{x}_t^{\text{det}}$ denote solutions of the deterministic equation

$$\varepsilon \frac{\mathrm{d}x}{\mathrm{d}t} = f(x, t) \tag{4.19}$$

tracking, respectively, the stable equilibrium branch $x_+^\star(t)$ and the unstable equilibrium branch $x_0^\star(t)$, while x_t denotes a general solution of the SDE (4.1). Our aim is to establish an upper bound for the probability *not* to reach a level δ_0 between $x_0^\star(t)$ and $x_-^\star(t)$, situated at a distance of order 1 from both equilibria. [BG, Theorem 2.3] shows that if x_t reaches δ_0, and δ_0 is close enough to $x_-^\star(t)$ (but it may still be at a distance of order 1), then it is likely to reach a neighbourhood of $x_-^\star(t)$ as well.

Let $\delta_0 < \delta_1 < x_{\text{c}} < \delta_2$ be the constants satisfying (2.46), that is,

$$\begin{aligned} f(x,t) &\asymp -1 && \text{for } \delta_0 \leqslant x \leqslant \delta_1 \text{ and } |t| \leqslant T \\ \partial_{xx} f(x,t) &\leqslant 0 && \text{for } \delta_1 \leqslant x \leqslant \delta_2 \text{ and } |t| \leqslant T. \end{aligned} \tag{4.20}$$

The basic ingredient of our estimate is the following analogue of Lemma 3.11:

Lemma 4.4. *Fix some initial time $t_0 \in [-T, T]$. We consider the following processes on $[t_0, T]$:*

- *the solution x_t^{det} of the deterministic differential equation (4.19) with an initial condition $x_{t_0}^{\text{det}} \in [\delta_1, \delta_2]$, such that $x_t^{\text{det}} \geqslant \delta_1$ for all $t \in [t_0, T]$;*
- *the solution x_t of the SDE (4.1) with an initial condition $x_{t_0} \in [x_{t_0}^{\text{det}}, \delta_2]$;*
- *the difference $y_t = x_t - x_t^{\text{det}}$, which satisfies $y_{t_0} = x_{t_0} - x_{t_0}^{\text{det}} \geqslant 0$;*

- *the solution y^0_t of the linear SDE*

$$\mathrm{d}y^0_t = \frac{1}{\varepsilon}\tilde{a}(t)y^0_t\,\mathrm{d}t + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_t, \qquad \textit{where } \tilde{a}(t) = \partial_x f(x^{\det}_t, t) \tag{4.21}$$

  with initial condition $y^0_{t_0} \in [y_{t_0}, \delta_2 - x^{\det}_{t_0}]$.

If $\delta_1 \leqslant y^0_s + x^{\det}_s \leqslant \delta_2$ for all $s \in [t_0, t]$, then $y_s \leqslant y^0_s$ for those s. Similarly, if $\delta_1 \leqslant x_s \leqslant \delta_2$ for all $s \in [t_0, t]$, then $y^0_s \geqslant y_s$ for those s. The result remains true when t is replaced by a stopping time.

We will proceed as follows. Let $\zeta(t)$ be the function defined in (4.18), and let h be such that $x^{\det}_s + h\sqrt{\zeta(s)} < \delta_2$ for all $s \in [t_0, t]$. Given $x_0 \in (\delta_1, \delta_2)$ and times $t_0 < t_1 < t$ in $[-T, T]$, we consider the solution x_t of the SDE (4.1) with initial condition $x_{t_0} = x_0$. We introduce the stopping time

$$\tau = \inf\bigl\{s \in [t_0, t_1] \colon x_s \leqslant \delta_1\bigr\} \in [t_0, t_1] \cup \{\infty\}. \tag{4.22}$$

We can thus write

$$\begin{aligned}\mathbb{P}^{t_0,x_0}\bigl\{x_s > \delta_0 \;\; \forall s \in [t_0, t]\bigr\} \leqslant{}& \mathbb{P}^{t_0,x_0}\bigl\{x_s > \delta_1 \;\; \forall s \in [t_0, t_1]\bigr\} \\ &+ \mathbb{E}^{t_0,x_0}\Bigl\{1_{\{\tau \leqslant t_1\}}\mathbb{P}^{\tau,\delta_1}\bigl\{x_s > \delta_0 \;\; \forall s \in [\tau, t]\bigr\}\Bigr\}.\end{aligned} \tag{4.23}$$

The first term can be further estimated by

$$\begin{aligned}\mathbb{P}^{t_0,x_0}\bigl\{x_s > \delta_1 \;\; \forall s \in [t_0, t_1]\bigr\} \leqslant{}& \mathbb{P}^{t_0,x_0}\Bigl\{\sup_{t_0 \leqslant s \leqslant t_1} \frac{x_s - x^{\det}_s}{\sqrt{\zeta(s)}} > h\Bigr\} \\ &+ \mathbb{P}^{t_0,x_0}\bigl\{\delta_1 < x_s \leqslant x^{\det}_s + h\sqrt{\zeta(s)} \;\; \forall s \in [t_0, t_1]\bigr\}.\end{aligned} \tag{4.24}$$

The two summands in (4.24) can be estimated in a similar way as in the symmetric case. The first one is dealt with in the following result.

Proposition 4.5. *Assume $\delta_1 \leqslant x_0 \leqslant x^{\det}_{t_0} + \frac{1}{2}h\sqrt{\zeta(t_0)}$. Then*

$$\mathbb{P}^{t_0,x_0}\Bigl\{\sup_{t_0 \leqslant s \leqslant t_1} \frac{x_s - x^{\det}_s}{\sqrt{\zeta(s)}} > h\Bigr\} \leqslant \frac{5}{2}\Bigl(\frac{|\overline{\alpha}(t_1, t_0)|}{\varepsilon} + 1\Bigr)\,\mathrm{e}^{-\kappa h^2/\sigma^2}, \tag{4.25}$$

where κ is a positive constant and $\overline{\alpha}(t_1, t_0) = \int_{t_0}^{t_1} \bar{a}(s)\,\mathrm{d}s$.

The proof is almost the same as the proof of Proposition 3.12. Instead of (3.84), we may define τ^0 and τ^+ as the first times when $y^{(k)}_s$ either reaches $\delta_1 - x^{\det}_s$ or the upper boundary $h\sqrt{\zeta(s)}$. Then Lemma 4.4 implies $y_s \leqslant y^{(k)}_s$ for $s \in [u_k, u_{k+1} \wedge \tau^0 \wedge \tau^+]$. However, when estimating the probability that $\tau^0 < u_{k+1}$ as in (3.88), it is sufficient to use the fact that τ^0 is larger than the first time $y^{(k)}_s$ reaches 0. Finally, (3.91) still holds with the present definitions of ζ and $\bar{a}$, because of (4.18) and Proposition 4.3.

Let us now examine the second term in (4.24).

Proposition 4.6. *There exist constants $c_1 > 0$ and $\bar\kappa > 0$ such that, if $c_1^{3/2}\sigma \geqslant a_0^{3/4} \vee \varepsilon^{3/4}$ and $h > 2\sigma$, then*

$$\begin{aligned}\mathbb{P}^{-c_1\sigma^{2/3},x_0}&\bigl\{\delta_1 < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [-c_1\sigma^{2/3}, t_1]\bigr\} \\ &\leqslant \frac{3}{2}\exp\biggl\{-\bar\kappa\,\frac{1}{\log(h/\sigma)\vee|\log\sigma|}\,\frac{\widehat\alpha(t_1,-c_1\sigma^{2/3})}{\varepsilon}\biggr\}\end{aligned} \tag{4.26}$$

holds with $\widehat\alpha(t,s) = \int_s^t \widehat a(u)\,\mathrm{d}u$, for $-c_1\sigma^{2/3} \leqslant t_1 \leqslant c_1\sigma^{2/3}$ and all initial conditions x_0 satisfying $\delta_1 \leqslant x_0 \leqslant x^{\det}_{-c_1\sigma^{2/3}} + h\sqrt{\zeta(-c_1\sigma^{2/3})}$.

Proof:

1. Let $\widehat x_t^{\det}$ be the deterministic solution tracking the saddle at $x_0^\star(t)$ and set $x_t = \widehat x_t^{\det} + z_t$. Then

$$\mathrm{d}z_t = \frac{1}{\varepsilon}\bigl[\widehat a(t)z_t + \widehat b(z_t,t)\bigr]\,\mathrm{d}t + \frac{\sigma}{\sqrt\varepsilon}\,\mathrm{d}W_t, \tag{4.27}$$

where (4.17), (4.18) and (4.20) imply

$$\widehat a(t) \asymp |t| \vee \sqrt{a_0} \vee \sqrt\varepsilon \asymp \frac{1}{\zeta(t)} \tag{4.28}$$

and

$$\widehat b(z,t) \leqslant 0 \qquad\qquad \text{for } \widehat x_t^{\det} + z \in [\delta_1, \delta_2]. \tag{4.29}$$

Let $\varrho = \varrho(h/\sigma) \geqslant 1$ and define a partition $-c_1\sigma^{2/3} = u_0 < \dots < u_K = t_1$ of $[-c_1\sigma^{2/3}, t]$ by

$$\widehat\alpha(u_k, u_{k-1}) = \varrho\varepsilon \qquad \text{for } 1 \leqslant k < K = \biggl\lceil \frac{\widehat\alpha(t_1, -c_1\sigma^{2/3})}{\varrho\varepsilon}\biggr\rceil. \tag{4.30}$$

Let

$$Q_k = \sup_{\delta_1 < \widehat x_{u_k}^{\det} + z_k \leqslant x_{u_k}^{\det} + h\sqrt{\zeta(u_k)}} \mathbb{P}^{u_k,z_k}\bigl\{\delta_1 < \widehat x_s^{\det} + z_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, u_{k+1}]\bigr\}. \tag{4.31}$$

Then we have, as in (3.99),

$$\mathbb{P}^{-c_1\sigma^{2/3},x_0}\bigl\{\delta_1 < x_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [-c_1\sigma^{2/3}, t_1]\bigr\} \leqslant \prod_{k=0}^{K-1} Q_k. \tag{4.32}$$

The result will thus be proved if we manage to choose ϱ in such a way that Q_k is bounded away from 1 for $k = 0, \dots, K-2$.

2. We will estimate the Q_k in a similar way as in Proposition 3.13, but we shall distinguish three cases instead of two. These cases correspond to x_s reaching the levels $x_s^{\det}$, $\widehat{x}_s^{\det}$ and δ_1. We introduce a subdivision $u_k < \tilde{u}_{k,1} < \tilde{u}_{k,2} < u_{k+1}$ defined by

$$\widehat{\alpha}(\tilde{u}_{k,1}, u_k) = \frac{1}{3}\varrho\varepsilon, \qquad \widehat{\alpha}(\tilde{u}_{k,2}, u_k) = \frac{2}{3}\varrho\varepsilon, \tag{4.33}$$

and stopping times

$$\begin{aligned}\tau_{k,1} &= \inf\bigl\{s \in [u_k, \tilde{u}_{k,1}] \colon z_s \leqslant x_s^{\det} - \widehat{x}_s^{\det}\bigr\} \\ \tau_{k,2} &= \inf\bigl\{s \in [u_k, \tilde{u}_{k,2}] \colon z_s \leqslant 0\bigr\}.\end{aligned} \tag{4.34}$$

Then we can write, similarly as in (3.110),

$$\begin{aligned}&\mathbb{P}^{u_k,z_k}\bigl\{\delta_1 < \widehat{x}_s^{\det} + z_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, u_{k+1}]\bigr\} \qquad (4.35)\\ &\leqslant \mathbb{P}^{u_k,z_k}\bigl\{x_s^{\det} < \widehat{x}_s^{\det} + z_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [u_k, \tilde{u}_{k,1}]\bigr\} \\ &\quad + \mathbb{E}^{u_k,z_k}\Bigl\{1_{\{\tau_{k,1} < \tilde{u}_{k,1}\}}\mathbb{P}^{\tau_{k,1},z_{\tau_{k,1}}}\bigl\{\delta_1 < \widehat{x}_s^{\det} + z_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [\tau_{k,1}, u_{k+1}]\bigr\}\Bigr\}\end{aligned}$$

The first term can be bounded by comparing with the solution of the SDE (4.1) linearized around $x_t^{\det}$, with the help of Lemma 4.4. As in (3.112), we obtain the upper bound

$$\frac{2}{\sqrt{\pi}}\sqrt{\frac{1}{1-\mathrm{e}^{-2\varrho/3L}}} \sup_{u_k \leqslant u \leqslant \tilde{u}_{k,1}} \sqrt{|\bar{a}(u)|\zeta(u_k)}\,\frac{h}{\sigma}\,\mathrm{e}^{-\varrho/3L}, \tag{4.36}$$

where $L > 0$ is a constant such that $\widehat{a}(u) \leqslant L|\bar{a}(u)|$ for all u. Now if $\tau_{k,1} < \tilde{u}_{k,1}$, we also have

$$\begin{aligned}&\mathbb{P}^{\tau_{k,1},z_{\tau_{k,1}}}\bigl\{\delta_1 < \widehat{x}_s^{\det} + z_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [\tau_{k,1}, u_{k+1}]\bigr\} \\ &\quad\leqslant \mathbb{P}^{\tau_{k,1},z_{\tau_{k,1}}}\bigl\{0 < z_s \leqslant x_s^{\det} - \widehat{x}_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [\tau_{k,1}, \tilde{u}_{k,2}]\bigr\} \qquad (4.37)\\ &\quad\quad + \mathbb{E}^{\tau_{k,1},z_{\tau_{k,1}}}\Bigl\{1_{\{\tau_{k,2} < \tilde{u}_{k,2}\}} \\ &\qquad\qquad \mathbb{P}^{\tau_{k,2},z_{\tau_{k,2}}}\bigl\{\delta_1 < \widehat{x}_s^{\det} + z_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)} \;\; \forall s \in [\tau_{k,2}, u_{k+1}]\bigr\}\Bigr\}.\end{aligned}$$

Comparing with the solution of the SDE (4.1) linearized around $\widehat{x}_t^{\det}$, the first term can be bounded, as in (3.104), by

$$\frac{2}{\sqrt{\pi}}\sqrt{\frac{1}{1-\mathrm{e}^{-2\varrho/3}}} \sup_{u_k \leqslant u \leqslant \tilde{u}_{k,2}} \frac{\sqrt{\widehat{a}(u)}}{\sigma} \sup_{u_k \leqslant u \leqslant \tilde{u}_{k,2}} (x_u^{\det} - \widehat{x}_u^{\det}). \tag{4.38}$$

This estimate shows that a path starting on $x^{\det}$ at time $\tau_{k,1}$ has an appreciable probability to reach the saddle before time $\tilde{u}_{k,2}$. Note, however, that we cannot obtain directly a similar estimate for the probability to reach δ_1 as well, which is why we restart the process in $\widehat{x}^{\det}$.

3. In order to estimate the second summand in (4.37), let $z_s^{(k)}$ be the process starting in 0 at time $\tau_{k,2}$ which satisfies the linear equation

$$\mathrm{d}z_s^{(k)} = \frac{1}{\varepsilon}\widehat{a}(s)z_s^{(k)}\,\mathrm{d}s + \frac{\sigma}{\sqrt{\varepsilon}}\,\mathrm{d}W_s^{(k)}, \tag{4.39}$$

with $W_s^{(k)} = W_s - W_{\tau_{k,2}}$. The variance $v_{u_{k+1}}^{(k)}$ of $z_{u_{k+1}}^{(k)}$ satisfies, as in (3.101),

$$\mathrm{e}^{-2\widehat{\alpha}(u_{k+1},\tau_{k,2})/\varepsilon}\, v_{u_{k+1}}^{(k)} \geqslant \frac{1-\mathrm{e}^{-2\varrho/3}}{2} \inf_{u_k \leqslant u \leqslant u_{k+1}} \frac{\sigma^2}{\widehat{a}(u)}. \tag{4.40}$$

Thus we obtain, using Lemma 4.4,

$$\begin{aligned}
&\mathbb{P}^{\tau_{k,2},z_{\tau_{k,2}}}\Bigl\{\delta_1 < \widehat{x}_s^{\det} + z_s \leqslant x_s^{\det} + h\sqrt{\zeta(s)}\ \ \forall s \in [\tau_{k,2}, u_{k+1}]\Bigr\} \\
&\quad\leqslant \mathbb{P}^{\tau_{k,2},0}\bigl\{z_{u_{k+1}}^{(k)} \geqslant -(\widehat{x}_{u_{k+1}}^{\det} - \delta_1)\bigr\} \\
&\quad= \frac{1}{\sqrt{2\pi}} \int_{-(\widehat{x}_{u_{k+1}}^{\det}-\delta_1)(v_{u_{k+1}}^{(k)})^{-1/2}}^{\infty} \mathrm{e}^{-z^2/2}\,\mathrm{d}z \\
&\quad\leqslant \frac{1}{2} + \frac{1}{\sqrt{\pi}}\sqrt{\frac{1}{1-\mathrm{e}^{-2\varrho/3}}} \sup_{u_k \leqslant u \leqslant u_{k+1}} \frac{\sqrt{\widehat{a}(u)}}{\sigma} (\widehat{x}_{u_{k+1}}^{\det} - \delta_1)\,\mathrm{e}^{-\varrho/3}\,.
\end{aligned} \tag{4.41}$$

Here the introduction of the stopping time $\tau_{k,2}$ turns out to play a crucial role. The above probability is indeed close to $1/2$ when ϱ is larger than a constant times $|\log\sigma|$, which shows that once a path has reached the saddle, it also has about fifty percent chance to reach the level δ_1 in a time of order $\varepsilon|\log\sigma|/\widehat{a}(u)$.

4. From (4.36), (4.38) and (4.41) and the fact that $\varrho \geqslant 1$, we obtain the existence of a constant $C > 0$ such that

$$\begin{aligned}
Q_k \leqslant \frac{1}{2} + \frac{C}{\sigma} \sup_{u_k \leqslant u \leqslant u_{k+1}} \sqrt{\widehat{a}(u)}\Bigl[\sqrt{\zeta(u_k)}\,h\,\mathrm{e}^{-\varrho/3L} \\
+ \sup_{u_k \leqslant u \leqslant u_{k+1}} (x_u^{\det} - \widehat{x}_u^{\det}) + (\widehat{x}_{u_{k+1}}^{\det} - \delta_1)\,\mathrm{e}^{-\varrho/3}\Bigr].
\end{aligned} \tag{4.42}$$

Since $|u| \leqslant c_1\sigma^{2/3}$, the properties of $x^{\det}$, $\widehat{x}^{\det}$, ζ and $\widehat{a}$ imply the existence of another constant C_1 such that

$$Q_k \leqslant \frac{1}{2} + C_1\Bigl[\frac{h}{\sigma}\,\mathrm{e}^{-\varrho/3L} + c_1^{3/2} + \sqrt{c_1}\,\frac{\mathrm{e}^{-\varrho/3}}{\sigma^{2/3}}\Bigr]. \tag{4.43}$$

Now choosing $c_1^{3/2} = 1/(18C_1)$ and $\varrho = 3L\log(18C_1h/\sigma) \vee 3\log(18C_1\sqrt{c_1}/\sigma^{2/3}) \vee 1$, we get $Q_k \leqslant 2/3$ for $k = 0,\dots,K-2$ and thus

$$\prod_{k=0}^{K-1} Q_k \leqslant \frac{3}{2}\frac{1}{(3/2)^K} \leqslant \frac{3}{2}\exp\Bigl\{-\Bigl(\log\frac{3}{2}\Bigr)\frac{\widehat{\alpha}(t_1,-c_1\sigma^{2/3})}{\varrho\varepsilon}\Bigr\}, \tag{4.44}$$

and the result follows from our choice of ϱ. □

It remains to estimate the second term in (4.23), describing the probability not to reach δ_0 when starting in δ_1. This is done by using the fact that, by assumption, the drift term is bounded away from zero on the interval $[\delta_0, \delta_1]$. We will need to assume that it can be bounded away from zero on a slightly larger interval, which is possible by continuity of f.

Proposition 4.7. *Let $0 < \rho \leqslant \delta_1 - \delta_0$ be a constant such that $f(x,t) \leqslant -f_0 < 0$ for $\delta_0 \leqslant x \leqslant \delta_1 + \rho$ and $|t| \leqslant T$. Then*

$$\mathbb{P}^{t_0,\delta_1}\bigl\{x_s > \delta_0\ \ \forall s \in [t_0, t_0 + c\varepsilon]\bigr\} \leqslant \mathrm{e}^{-\tilde{\kappa}/\sigma^2} \tag{4.45}$$

holds for all $t_0 \in [-T, T - c\varepsilon]$, where $\tilde{\kappa} = f_0\rho^2/4(\delta_1 - \delta_0)$ and $c = 2(\delta_1 - \delta_0)/f_0$.

Proof: Let x^0_t be defined by

$$x^0_t = \delta_1 - \frac{f_0}{\varepsilon}(t-t_0) + \frac{\sigma}{\sqrt{\varepsilon}} W_{t-t_0}, \qquad t \geqslant t_0. \tag{4.46}$$

By Gronwall's inequality, it is easy to see, as in Lemma 4.4, that if $\delta_0 \leqslant x^0_t \leqslant \delta_1 + \rho$ for all $t \in [t_0, t_0 + c\varepsilon]$, then $x_t \leqslant x^0_t$ for those t. We thus have

$$\begin{aligned} \mathbb{P}^{t_0,\delta_1}\bigl\{x_s > \delta_0 \;\; \forall s \in [t_0, t_0+c\varepsilon]\bigr\} &\leqslant \mathbb{P}^{t_0,\delta_1}\Bigl\{\sup_{t_0 \leqslant s \leqslant t_0 + c\varepsilon} x^0_s + \frac{f_0}{\varepsilon}(s-t_0) > \delta_1 + \rho\Bigr\} \\ &\quad + \mathbb{P}^{t_0,\delta_1}\Bigl\{\delta_0 < x^0_s < \delta_1 + \rho - \frac{f_0}{\varepsilon}(s-t_0) \;\; \forall s \in [t_0, t_0 + c\varepsilon]\Bigr\}. \end{aligned} \tag{4.47}$$

Note, however, that for $s = t_0 + c\varepsilon$,

$$\delta_1 + \rho - \frac{f_0}{\varepsilon}(s - t_0) = \delta_1 + \rho - 2(\delta_1 - \delta_0) \leqslant \delta_0, \tag{4.48}$$

so that the second term in (4.47) is equal to zero. The first term equals

$$\mathbb{P}^{0,0}\Bigl\{\sup_{0 \leqslant s \leqslant c\varepsilon} \frac{\sigma}{\sqrt{\varepsilon}} W_s > \rho\Bigr\} \leqslant \exp\Bigl\{-\frac{\rho^2}{2c\sigma^2}\Bigr\} \tag{4.49}$$

by Doob's submartingale inequality. □

References

[BPSV] R. Benzi, A. Sutera, A. Vulpiani, *The mechanism of stochastic resonance*, J. Phys. A **14**:L453–L457 (1981).

R. Benzi, G. Parisi, A. Sutera, A. Vulpiani, *A theory of stochastic resonance in climatic change*, SIAM J. Appl. Math. **43**:565–578 (1983).

[BG] N. Berglund, B. Gentz, *Pathwise description of dynamic pitchfork bifurcations with additive noise*, WIAS preprint no. 604 (2000).
`http://www.wias-berlin.de/publications/preprints/604`
`arXiv:math.PR/0008208`

[ET] J.-P. Eckmann, L. E. Thomas, *Remarks on stochastic resonance*, J. Phys. A **15**:L261–L266 (1982).

[FH] S. Fauve, F. Heslot, *Stochastic resonance in a bistable system*, Phys. Lett. **97A**:5–7 (1983).

[Fox] R. F. Fox, *Stochastic resonance in a double well*, Phys. Rev. A **39**:4148–4153 (1989).

[Fr] M. I. Freidlin, *Quasi-deterministic approximation, metastability and stochastic resonance*, Physica D **137**:333–352 (2000).

[Gr] I. S. Gradšteĭn, *Applications of A. M. Lyapunov's theory of stability to the theory of differential equations with small coefficients in the derivatives*, Mat. Sbornik N.S. **32**:263–286 (1953).

[JH] P. Jung, P. Hänggi, *Stochastic nonlinear dynamics modulated by external periodic forces*, Europhys. Letters **8**:505–510 (1989).

P. Jung, P. Hänggi, *Amplification of small signals via stochastic resonance*, Phys. Rev. A **44**:8032–8042 (1991).

[MW] F. Moss, K. Wiesenfeld, *The benefits of background noise*, Scientific American **273**:50–53 (1995).

K. Wiesenfeld, F. Moss, *Stochastic resonance and the benefits of noise: from ice ages to crayfish and SQUIDs*, Nature **373**:33–36 (1995).

[McNW] B. McNamara, K. Wiesenfeld, *Theory of stochastic resonance*, Phys. Rev. A **39**:4854–4869 (1989).

[SNAS] B. Shulgin, A. Neiman, V. Anishchenko, *Mean switching frequency locking in stochastic bistable systems driven by a periodic force*, Phys. Rev. Letters **75**:4157–4160 (1995).

A. Neiman, A. Silchenko, V. Anishchenko, L. Schimansky-Geier, *Stochastic resonance: Noise-enhanced phase coherence*, Phys. Rev. E **58**:7118–7125 (1998).

[Ti] A. N. Tihonov, *Systems of differential equations containing small parameters in the derivatives*, Mat. Sbornik N.S. **31**:575–586 (1952).

Nils Berglund
Department of Mathematics, ETH Zürich
ETH Zentrum, 8092 Zürich, Switzerland
E-mail address: berglund@math.ethz.ch

Barbara Gentz
Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstraße 39, 10117 Berlin, Germany
E-mail address: gentz@wias-berlin.de